

**Laëtitia GUARDINI**

M2P Bioingénierie-biotechnologies végétales

**Année 2011-2012**

Tuteur : Christophe Jacquet

# **Rôle des TALEs dans la spécialisation du pouvoir pathogène de *Xanthomonas citri* pv. *citri***

Aline ESCALON ; Lionel GAGNEVIN

Equipe Génomique et épidémiologie des agents pathogènes  
émergents

Responsable : Olivier PRUVOST

UMR PVBMT, pôle de protection des plantes, CIRAD  
7 ch de l'Irat 97410 St Pierre, Ile de la Réunion



Université  
Paul Sabatier  
TOULOUSE III



## **REMERCIEMENTS**

Je tiens en premier lieu à remercier Aline Escalon, ma maitre de stage, pour son enthousiasme et sa bonne humeur, les connaissances qu'elle m'a apportées et a su partager tout au long du projet. Elle m'a permis de conserver une véritable autonomie en restant toujours présente pour me conseiller et m'épauler. Ces conditions travail idéales ont rendu ce stage particulièrement enrichissant et formateur. .

Je tiens également à remercier Lionel Gagnevin, mon maitre de stage, pour son accompagnement et ses conseils. Son expérience et son savoir faire ont permis de débloquer des situations complexes et permettre ainsi la bonne continuité du projet.

J'exprime toute ma reconnaissance à Olivier Pruvost, pour m'avoir accueillie dans son équipe et pour ses nombreux conseils avisés. Ainsi qu'à Christian Vernière pour ses connaissances et son soutien.

Je tiens à dire un grand merci à Stephanie Javegny pour son aide précieuse, sa technicité et son savoir faire.

Je tiens enfin à remercier toutes les personnes du pôle de protection des plantes, l'équipe de bio-mol et l'ensemble des stagiaires, VSCs et thésards avec qui j'ai pu partager le quotidien, travailler et enrichir mes connaissances en découvrant de nouveaux domaines scientifiques.



## **ABBREVIATIONS**

**ADN:** Acides Desoxy ribonucléotides

**BET:** Bromure éthidium

**BSA:** Serum Albumine Bovine

**dNTP:** Desoxyribo Nucléotide Tri Phosphate

***E. coli:*** *Escherichia coli*

**EDTA:** acide éthylène diamine tétraacétique

**FLK:** Région flanquante

**IPTG:** IsoPropyl  $\beta$ -1-D ThioGalactopyranoside

**LB:** Luria-Bertani

**LPGA:** Levure Peptone Glucose Agar

**PBS:** Phosphate Buffered Saline (Tampon phosphate salin )

**PCR:** Polymerase Chain Reaction

**pv. :** Pathovar

**RVD:** repeat variable diresidue

**SSTT:** Système de Sécrétion de Type Trois

**TAE:** Tris Acetate EDTA (tampon)

**TALE:** Transcription activator Like Effector (effecteur mimant les facteurs de transcription)

***Xcc:*** *Xanthomonas citri* pathovar. *citri*

**X-gal:** 5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside



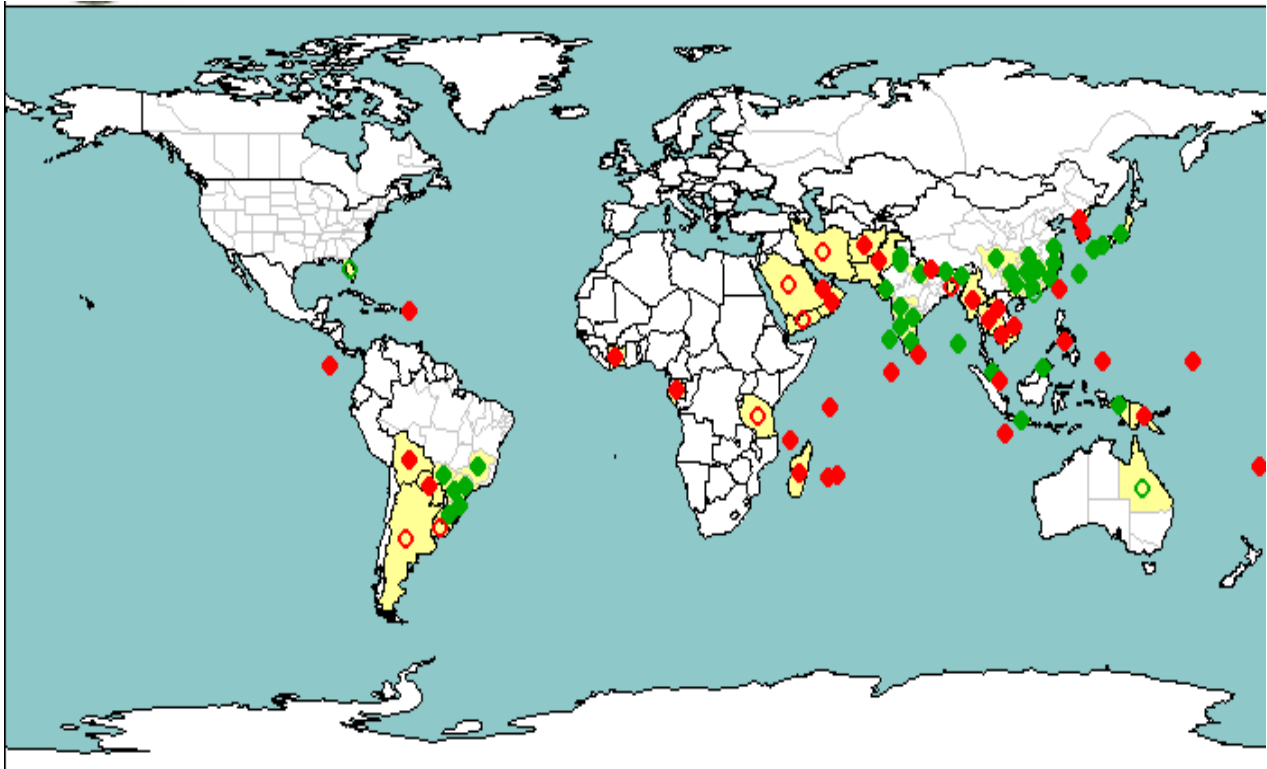
## Sommaire

1	Introduction.....	11
1.1	Présentation du CIRAD.....	11
1.2	Le chancre asiatique des agrumes .....	11
1.2.1	Origine et répartition de la maladie.....	11
1.2.2	Les symptômes de la maladie.....	13
1.2.3	Le cycle de la maladie .....	13
1.3	<i>Xanthomonas citri</i> pv. <i>citri</i> ( <i>Xcc</i> ).....	15
1.3.1	Identification et caractéristiques.....	15
1.3.2	Le génome de <i>Xanthomonas citri</i> pv. <i>citri</i> .....	17
1.3.3	Le pouvoir pathogène de <i>Xcc</i> .....	17
1.4	Les TALEs (Transcription Activator Like Effectors).....	21
1.4.1	Structure et interaction avec l'ADN végétal .....	21
1.4.2	Rôles mis en évidence .....	21
1.4.2.1	Rôle dans la formation du chancre des agrumes.....	23
1.4.2.2	Rôle dans la restriction de la gamme d'hôte .....	23
1.4.3	Diversité des TALEs chez <i>Xcc</i> .....	25
1.5	Projet et enjeux.....	27
2	Matériel et méthodes .....	31
2.1	Matériel .....	31
2.1.1	Agrumes .....	31
2.1.2	Souches bactériennes et vecteurs .....	31
2.2	Méthodes .....	33
2.2.1	Amplification et clonage des TALEs .....	33
2.2.1.1	Extraction d'ADN génomique .....	33
2.2.1.2	Dosage de l'ADN.....	33
2.2.1.3	PCR : Polymerase Chain Reaction .....	33
2.2.1.4	Electrophorèse de l'ADN sur gel d'agarose .....	35
2.2.1.5	Purification des amplicons.....	35





2.2.1.6	Clonage des TALEs dans pGEM®-T Easy .....	35
2.2.1.7	Electroperméabilisation d' <i>E.coli</i> TG1 .....	35
2.2.1.8	Transformation de bactéries compétentes et préparation plasmidique .....	37
2.2.1.9	Vérification de la qualité des extractions plasmidiques.....	37
2.2.1.10	Séquençage .....	39
2.2.2	Délétion du gène <i>pthA</i> *4 .....	39
2.2.2.1	Amplification des régions flanquantes et purification .....	39
2.2.2.2	Clonage Golden-gate .....	41
2.2.2.3	Conjugaison tri-parentale et sélection des mutants.....	43
2.2.2.4	Caractérisation phénotypique des mutants.....	45
2.2.2.5	Extraction sur symptômes et vérification des génotypes par PCR. ....	45
3	Résultats et discussion .....	47
3.1	Caractérisation des TALEs chez <i>Xanthomonas citri</i> pv. <i>citri</i> .....	47
3.1.1	Mise au point des différentes étapes .....	47
3.1.1.1	Amplification par PCR .....	49
3.1.1.2	Purification des produits PCR et ligation dans pGEM®-T Easy.....	49
3.1.1.3	Minipréparations de plasmides pour séquençage. ....	51
3.1.2	Résultats de séquençage .....	51
3.1.2.1	Recherche d'homologues et relevé des RVDs.....	51
3.1.2.2	Design d'amorces pour les répétitions centrales.....	53
3.1.2.3	Alignement des boxes promotrices .....	55
3.1.2.4	Recherche des promoteurs cibles, <i>in silico</i> .....	57
3.2	Délétion du gène <i>pthA</i> *4 chez 3 souches de <i>Xcc</i> A* .....	59
3.2.1	Amplification et clonage des régions flanquantes.....	59
3.2.2	Résultat de la conjugaison tri-parentale .....	61
3.2.3	Résultat de phénotypes sur plantes.....	61
3.2.4	Vérification des génotypes par PCR sur symptômes .....	63
4	Conclusion et perspectives:.....	65
	Bibliographie .....	69



**Figure 1 : Répartition du Chancre Asiatique des Agrumes dans le monde**

Sur cette carte est représentée la répartition mondiale du chancre asiatique des agrumes, à partir des données de l'EPPO de 2006. En rouge les données nationales et en vert les données sous-nationales. Les ronds pleins sont les régions où le chancre est présent sur toute la zone et les ronds vides celles où il n'est présent que dans certaines zones.

# **1 Introduction**

## **1.1 Présentation du CIRAD**

Le CIRAD (Centre de Coopération Internationale en Recherche Agronomique pour le Développement) a été créé en 1984, sous la forme d'un établissement public à caractère industriel et commercial. A la Réunion, le CIRAD est structuré en trois pôles de recherche dont le pôle de protection des plantes (3P). Ce pôle regroupe entre autre le CIRAD et l'Université de la Réunion formant l'Unité Mixte de Recherche « Peuplements Végétaux et Bioagresseurs en Milieu Tropical » (UMR PVBMT). Ses champs d'investigation sont l'épidémiologie végétale (épidémiosurveillance et adaptation des populations de micro-organismes phytopathogènes), la dynamique des populations d'insectes et les interactions tritrophiques, la génétique et la caractérisation de la résistance des plantes aux bioagresseurs, l'endémisme et les invasions des écosystèmes terrestres naturels en milieu insulaire.

L'UMR PVBMT est structuré en cinq équipes, dont celle responsable de cette étude : « Génomique et épidémiologie des agents pathogènes émergents ». Ses objectifs sont de mieux comprendre les facteurs et les mécanismes impliqués dans l'émergence des bactéries et virus phytopathogènes. Pour cela, sont analysées : les modifications des génomes et les forces évolutives associées (approche évolutive avec la mise en évidence des gènes associés au spectre d'hôtes de deux pathotypes de *Xanthomonas citri*, responsable du chancre asiatique des agrumes.) et les modifications des dynamiques épidémiologiques (approche épidémiologique dont l'analyse de la stabilité *in planta* de séquences-cibles chez *Xanthomonas citri* et la surveillance épidémiologique des populations de *Xanthomonas citri*.).

## **1.2 Le chancre asiatique des agrumes**

### **1.2.1 Origine et répartition de la maladie**

Le chancre asiatique des agrumes (CAA) également appelé chancre A ou vrai chancre, serait originaire d'Asie tropicale. Son origine exacte est encore un sujet de controverse (Das 2003) entre le sud-est de la Chine, l'Inde et l'Indonésie, zones d'où les agrumes hôtes seraient également originaires.

La maladie et l'agent causal ont été décrits aux USA en 1915 sur des plants de pépinière provenant d'Asie et plus précisément du Japon (Civerolo 1984). La maladie s'est rapidement propagée dans le monde au début du XX<sup>ème</sup> siècle, Amérique du sud (Rossetti 1977),



Afrique du sud (Doidge 1916) et Australie (Garnsey *et al.*, 1979). Des campagnes d'éradication ont été menées dans ces pays, par inspection des pépinières et vergers, mise en quarantaine et incinération des arbres infectés (Das 2003). Mais d'autres épidémies sont apparues, non seulement du sous continent Indien à l'est de l'Asie où la maladie est endémique, mais également dans d'autres régions telles-que les îles de l'Océan Indien, le Moyen-Orient, la péninsule Arabique, l'Afrique et de nouveau la Floride, l'Amérique du sud, l'Australie et les îles Fidji, entraînant des interdictions d'exportation responsables de fortes pertes économiques. Aujourd'hui seul le continent Européen (voir répartition figure 1) est encore épargné.

### 1.2.2 Les symptômes de la maladie

Le chancre Asiatique des agrumes, se caractérise par la formation de pustules chancreuses au niveau des tissus aériens, c'est-à-dire : tiges, feuilles et fruits (figure 2A).

Sur les feuilles, on observe dans un premier temps de petites taches aqueuses circulaires de 2-10 mm de diamètre, généralement sur la surface intérieure du limbe. Sur les feuilles, les tiges, les épines et les fruits, ces lésions circulaires se surélèvent et forment des cloques de plus en plus blanches, jusqu'à des pustules spongieuses jaunes. Ces pustules deviennent ensuite brun foncé, s'épaississent et sont rugueuses au toucher. Sur le long des tiges, des pustules peuvent confluer pour séparer l'épiderme. Les lésions sur les feuilles et les fruits matures sont parfois entourées d'un halo chlorotique jaune (qui peut disparaître) et un centre creux. La défoliation et la chute prématurée des fruits touchés se produisent sur des arbres fortement infectés (Brunings & Gabriel 2003).

L'observation histologique des symptômes du chancre asiatique des agrumes (figure 2A4), montre une hypertrophie (augmentation du volume cellulaire) et une hyperplasie cellulaire (augmentation du nombre de cellules) (Gürlebeck *et al.* 2006).

### 1.2.3 Le cycle de la maladie

L'infection des agrumes se fait au niveau des tissus jeunes en phase de croissance, qui sont les organes les plus sensibles (figure 2B). En effet à ce stade les stomates sont ouverts, sans que la cuticule soit totalement développée, permettant ainsi la pénétration d'inoculum (Graham *et al.* 2004). Chez les organes adultes à cuticule épaisse où l'eau s'infiltre plus difficilement et les organes très jeunes dépourvus de stomates fonctionnels, une blessure est



nécessaire pour l'infection. Ces blessures pouvant être provoquées par des vents violents, des outils agricoles, les épines des agrumes ou encore certains insectes.

La principale cause de dissémination de la maladie est la pluie associée au vent, sur des courtes et moyennes distances. Au sein des exploitations, la dissémination peut se faire également via les engins et les outils agricoles, ainsi que les vêtements de travail (figure 2C). Sur les longues distances la cause la plus probable reste la commercialisation et l'importation de plants contaminés (Gottwald *et al.* 2002).

### 1.3 *Xanthomonas citri* pv.*citri* (Xcc)

L'agent pathogène responsable du chancre asiatique des agrumes est la bactérie *Xanthomonas citri* pv. *citri* classée dans la division des Protéobactéries, classe des Gramma Protéobactéries, ordre des Xanthomonadales, famille des Xanthomonadaceae, genre *Xanthomonas*.

#### 1.3.1 Identification et caractéristiques

Entre la fin du 20<sup>ème</sup> et le début du 21<sup>ème</sup> siècle, de nombreuses études ont été menées sur la classification et la différenciation des espèces génomiques, des bactéries du genre *Xanthomonas*. Le pathovar *citri* a tout d'abord été classé au sein de l'espèce *Xanthomonas campestris*. Cette espèce a ensuite été reclassée en différentes espèces dont *Xanthomonas axonopodis*, où a été placé le pathovar *citri*. Enfin, grâce à des études récentes (Ah You *et al* 2009) le pathovar *citri* a été classé dans l'espèce *Xanthomonas citri*.

La classification des espèces en pathovars sépare les profils pathologiques des souches mais n'a aucune valeur taxonomique. Un pathovar regroupe des souches qui possèdent une gamme d'hôtes similaire et/ou qui induisent une même maladie. Le pathovar *citri* rassemble donc les bactéries de l'espèce *X.citri* responsables du chancre asiatique des agrumes. Pour définir la variabilité de gamme d'hôtes à des niveaux plus fins, c'est-à-dire au sein des agrumes pour *Xcc*, les bactéries d'un pathovar sont classifiées en pathotypes. On distingue chez *Xcc* trois pathotypes. Le pathotype A regroupe les souches à large gamme d'hôtes, c'est-à-dire pathogènes sur tous les agrumes décrits comme hôtes. Le pathotype A\* a une gamme d'hôtes restreinte à la lime mexicaine (*Citrus aurantifolia*) et à un porte greffe (*Citrus macrophylla*). Enfin le pathotype A<sup>w</sup> a été décrit (Sun *et al.* 2004). Les souches appartenant à ce pathotype





induisent une réponse d'hypersensibilité (HR) sur pomelos (*Citrus paradisi*), en plus de leur gamme d'hôtes identique à celle du pathotype A\* (restreinte à la lime mexicaine et au porte greffe *Citrus macrophylla*) Le pathotype A<sup>W</sup> peut donc être considéré comme un sous groupe du pathotype A\*.

*Xanthomonas citri* pv. *citri* est un bacille Gram-négatif, mobile par un flagelle polaire unique. La bactérie mesure 0.5-0.75 µm de large et 1.5-2 µm de long. Sur milieu nutritif gélosé, *Xcc* forme des colonies jaunes, circulaires, à contours réguliers, semi-bombées et brillantes après 3-4 jours d'incubation. *Xcc* est aérobie stricte, chimioorganotrophe avec un métabolisme oxydatif de glucose. C'est une bactérie psychrotrophe, dont la température optimale de croissance est de 28-30°C. Enfin elle produit un pigment jaune appelé la xanthomonadine et en présence de sucre elle produit un EPS (ExoPolySaccharide) très abondant rendant les colonies mucoïdes (Brunings & Gabriel 2003).

### 1.3.2 Le génome de *Xanthomonas citri* pv. *citri*

Le génome de la souche *Xcc* IAPAR 306 a été intégralement séquencé par une équipe brésilienne (da Silva *et al.* 2002). Ce séquençage a mis en évidence un chromosome circulaire de plus de 5 millions de paires de bases (5175554 pb) et deux plasmides : pXac33 (33699 pb) et pXac64 (64920 pb), portant tous deux les gènes *pthA* analysés dans cette étude.

La séquence de *Xcc* est un outil qui permet de nombreuses approches nouvelles et le développement d'outils pour l'étude des *Xanthomonas*, en particulier sur la relation hôte/pathogène et le pouvoir pathogène.

### 1.3.3 Le pouvoir pathogène de *Xcc*

Le pouvoir pathogène de *Xcc* est causé par de nombreux facteurs de virulence : les lipopolysaccharides (LPS) et exopolysaccharides (EPS), des adhésines ou des pili nécessaires à l'attachement des bactéries, les protéines de la membrane externe riches en glycine, ces dernières sont les principales structures de colonisation de l'hôte et de l'adhésion chez de nombreuses bactéries pathogènes.

La bactérie *Xanthomonas citri* pv. *citri*, possède plusieurs types de systèmes de sécrétion dont certains jouent un rôle dans le pouvoir pathogène. *Xcc* possède deux systèmes de sécrétion de type II, impliqués dans la sécrétion d'enzymes de dégradation de la paroi cellulaire, par exemple pour des enzymes cellulolytiques (Da-Silva *et al.* 2002) et de toxines ; deux

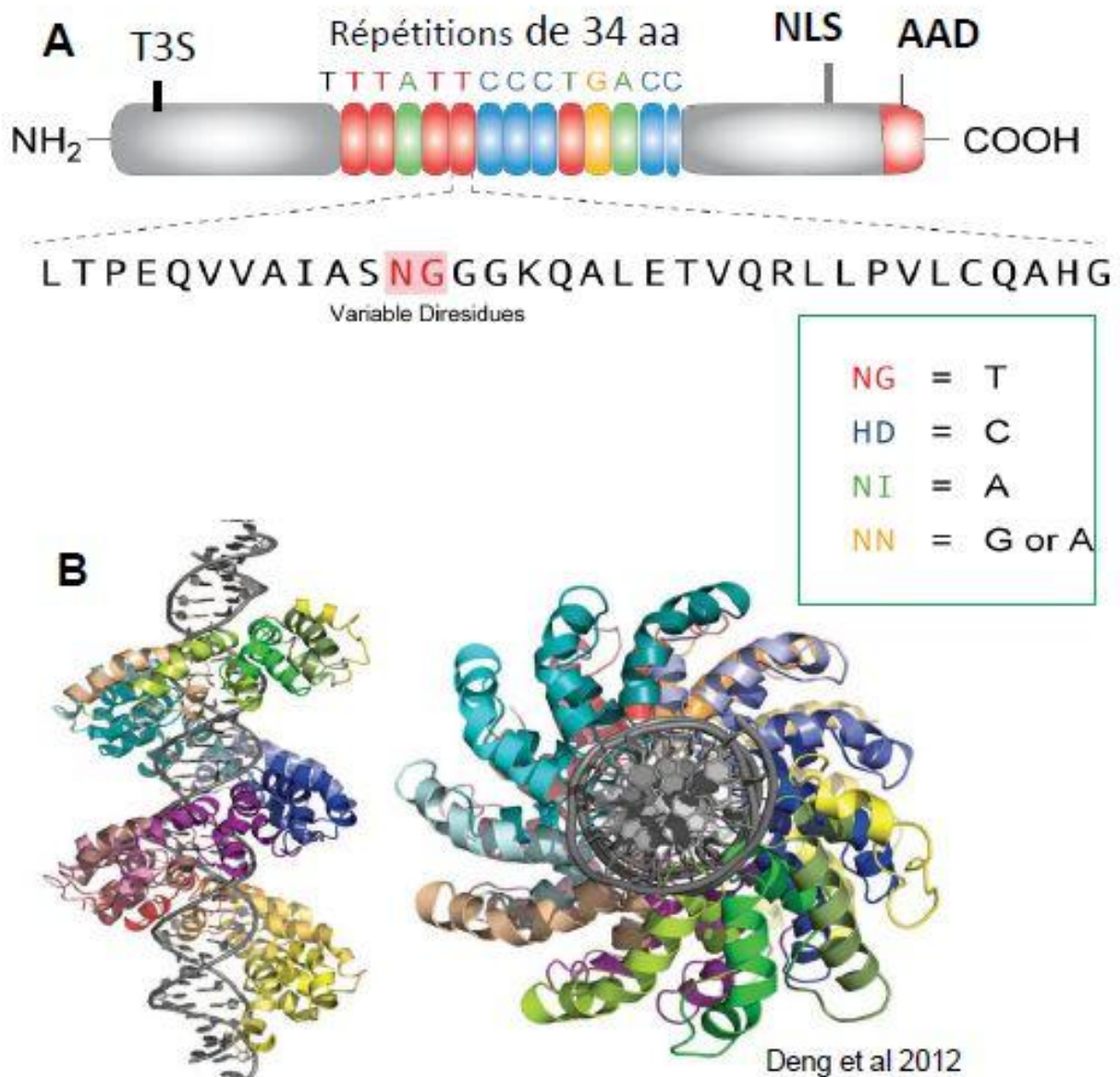


systèmes de sécrétion de type IV (T4S) codés par les gènes *vir* et impliqués dans le transport de protéines en général associées à de l'ADN (notamment au cours des phénomènes de conjugaison) et un système de sécrétion de type III (SSTT).

Principal acteur de la pathogénie de *Xcc*, le SSTT est un complexe protéique membranaire qui permet le transfert d'effecteurs de la bactérie à la cellule végétale. Le SSTT est codé par un ensemble de gène *hrp* (Hypersensitive Reaction and Pathogenicity) et joue un rôle à la fois dans le pouvoir pathogène sur un hôte sensible, ainsi que dans le déclenchement de la réaction d'hypersensibilité chez un hôte résistant.

La principale fonction du SSTT est l'injection de protéines dites « effectrices » (les effecteurs de type III) dans l'apoplaste et/ou dans le cytoplasme des cellules hôtes. Une souche de *Xcc* donnée est susceptible de sécréter via le SSTT, une trentaine de protéines appartenant à 60 familles et pouvant avoir différents rôles : dans l'avirulence en activant des gènes de résistance chez la plante et induisant ainsi une réaction d'hypersensibilité qui limite le spectre d'hôtes ; dans l'agressivité en contribuant au pouvoir pathogène (qualité et quantité de symptômes sur les plantes non résistantes) ou encore à la capacité adaptative de la bactérie. Une famille particulière existant chez de nombreux *Xanthomonas* est celle de AvrBs3/pthA (Avirulence/pathogénie), récemment rebaptisée famille des TALEs (Transcription Activator Like Effectors).

Chez *Xcc*, le gène *pth4* est directement impliqué dans la formation du chancre sur les agrumes (Brunings & Gabriel 2003). Tous les *Xcc* possèdent au moins deux homologues de *pthA4* (Lee, S *et al* 2008) le rôle de chacun de ces homologues n'étant toujours pas défini. Enfin, plusieurs homologues de la famille TALE ont été décrits chez d'autres espèces de *Xanthomonas*, par exemple les gènes *pthXo1*, *pthXo6* et *avrXa7* chez *Xanthomonas oryzae* pv. *oryzae*, l'agent pathogène responsable du flétrissement bactérien du riz (Römer *et al* 2010).



**Figure 3 : structure des TALEs**

A : structure schématique des TALEs comprenant: un domaine d'adressage au système de sécrétion de type III (T3S) en N-terminal ; un signal de localisation nucléaire (NLS) et un domaine d'activation (AAD) en C-terminal. La région centrale est composée de répétitions en tandem de 34 acides aminés, chacune de ces répétitions possédant en positions 12 et 13 deux acides aminés variables (Variable Diresidues), ces derniers reconnaissant spécifiquement une des bases du promoteur végétal. Dans l'encadré se trouve le code de liaison des dirésidus variables (RVD) les plus fréquemment retrouvés.

B : Structure de l'interaction TALE/ ADN obtenue par cristallographie. Le TALE se fixe au promoteur en s'insérant dans le grand sillon de l'ADN.

## 1.4 Les TALEs (Transcription Activator Like Effectors)

### 1.4.1 Structure et interaction avec l'ADN végétal

La structure des protéines TALEs est présentée figure 3A. Tous les TALEs possèdent un signal de sécrétion de type III en partie N-terminal et un signal de localisation nucléaire (NLS) ainsi qu'un domaine d'activation transcriptionnel eucaryote (AAD) en partie C-terminal. Les TALEs diffèrent entre eux au niveau de la région centrale. Celle-ci est constituée d'un nombre variable de répétitions de 33 à 35 acides aminés et d'une demi-répétition de 20 acides aminés coté C-terminal (Bogdanove *et al.* 2010). Le polymorphisme de ces répétitions se situe au niveau des acides aminés 12 et 13 qui forment un dirésidu variable au sein de chaque répétition (RVD).

Il a été démontré que c'est la région centrale des TALEs qui reconnaît et interagit avec l'ADN végétal, déterminant ainsi la spécificité de l'activation transcriptionnelle. Lors de l'interaction spécifique TALE/promoteur végétal, le TALE se fixe au promoteur en s'intercalant dans le grand sillon de l'ADN (Deng *et al.* 2012; Mak *et al.* 2012) (figure 3B).

Plusieurs travaux (Bogdanove *et al.* 2010 ;Boch *et al.* 2009) ont mis en évidence que la spécificité d'interaction entre la protéine et l'ADN se situe au niveau des RVDs, avec un "code de reconnaissance spécifique" des RVD les plus fréquents, pour une des bases azotée de l'ADN végétal. La combinaison d'acides aminés "HD" fixant une cytosine, "N\*" une thymine ou cytosine, "NG" une thymine, "NI" une adénine, le dirésidu "NS" étant aspécifique. Récemment, il a été montré que le RVD "NH" reconnaît une guanine (Streubel *et al.* 2012)

Un TALE fixe ainsi une ou plusieurs "boxes promotrices" (de tailles correspondantes au nombre de répétitions centrales du TALE), pour activer le ou les gènes situés en aval. S'il s'agit de gènes de susceptibilité (chez les plantes sensibles), il y aura formation de symptômes et s'il s'agit de gènes activant les défenses (sur plantes résistantes) il y aura résistance.

### 1.4.2 Rôles mis en évidence

De nombreux travaux ont cherché à caractériser le rôle des TALEs chez *Xcc* depuis que leur implication dans l'agressivité de cette bactérie a été démontrée (Swarup *et al.* 1991 ; Swarup *et al.* 1992).



### 1.4.2.1 Rôle dans la formation du chancre des agrumes

Une étude du rôle des TALEs dans la détermination de la gamme d'hôtes chez *Xcc* a été menée par Al-Saadi *et al* en 2007. Il a été mis en évidence que toutes les souches de *Xanthomonas citri* pv. *citri* possèdent en commun un TALE de 17.5 répétitions, et ce quelque soit le pathotype A, A\* ou A<sup>w</sup>. Afin de déterminer le rôle des différents TALEs dans la détermination de la gamme d'hôtes, le gène *pthA4* a été muté par insertion d'un transposon (Tn5), de façon à inactiver ce gène. Les mutants de chaque pathotype ont été inoculés sur plantes. Quelque soit le pathotype, l'inactivation du gène *pthA4* a conduit à la perte de la capacité à induire du chancre.

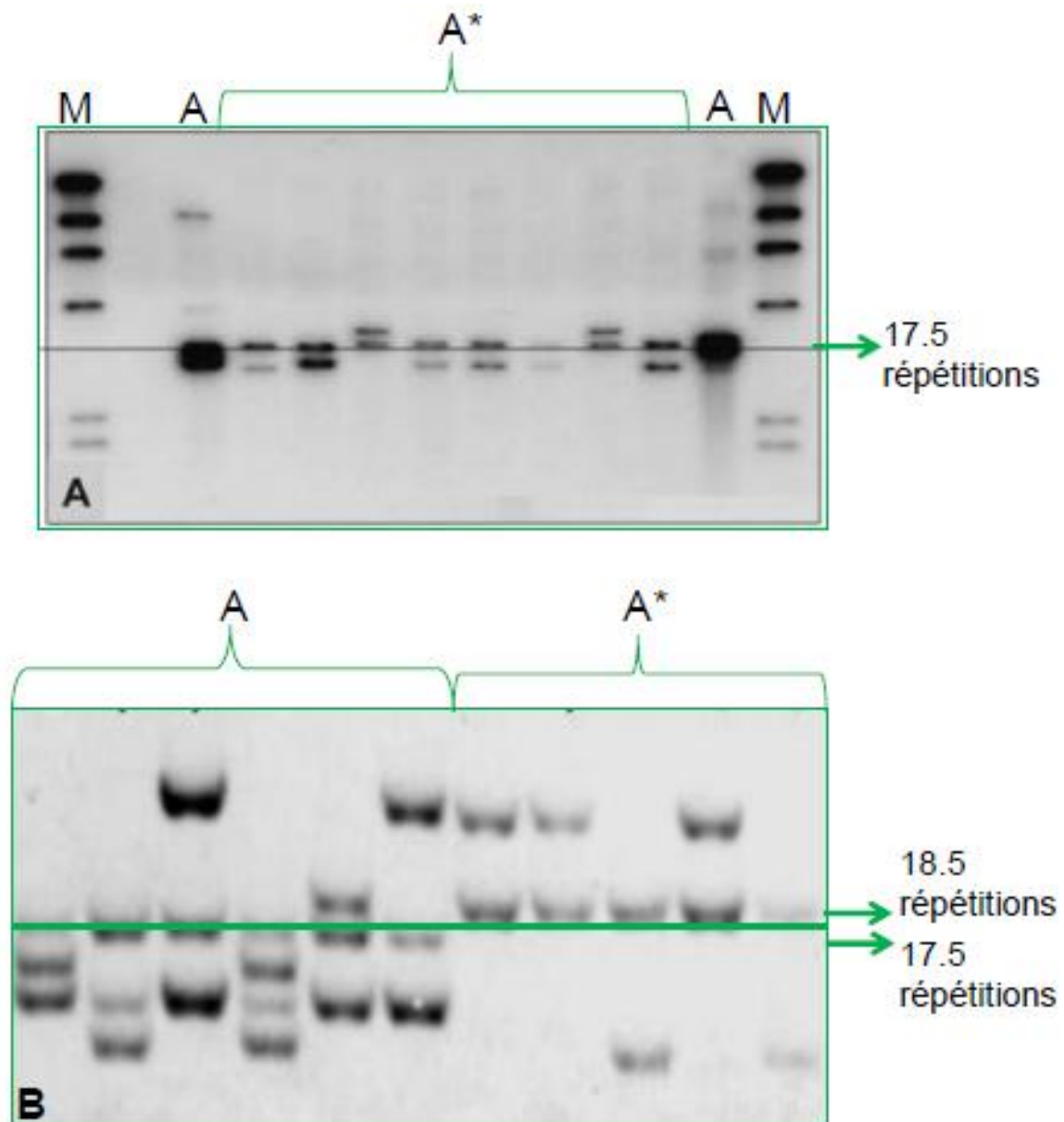
Afin de valider le rôle de *pthA4* dans la formation du chancre, des tests de complémentation ont été effectués sur les mutants à l'aide de deux plasmides contenant respectivement le gène étudié provenant d'une souche de pathotype A (*pthA4*) et d'une souche de pathotype A\* (*pthA\*4*). Quatre complémentations ont été testées : soucheAΔ*pthA4*::*pthA4* ; soucheAΔ*pthA4*::*pthA\*4* ; soucheA\*Δ*pthA\*4*::*pthA\*4* et soucheA\*Δ*pthA\*4*::*pthA4*. Les complémentations croisées permettant de voir si *pthA4* joue un rôle dans la spécialisation de la gamme d'hôtes.

La complémentation a permis de restaurer la capacité à faire du chancre pour toutes les souches, sans modification de la gamme d'hôtes. La souche du pathotype A conservant un large spectre d'hôtes malgré la présence de *pthA\*4* et inversement une souche A\* gardant une gamme d'hôtes restreinte avec la complémentation par *pthA4*. Finalement, cette étude a montré que le TALE *pthA4* est clairement impliqué dans la formation des symptômes de type chancreux, et n'intervient pas dans la détermination de la gamme d'hôtes entre les différents pathotypes de *Xcc*.

### 1.4.2.2 Rôle dans la restriction de la gamme d'hôte

Au cours de leurs investigations, Shiotani *et al* en 2007 ont remarqué que certaines souches de *Xcc* se différençaient en deux groupes distincts en fonction de leur agressivité forte ou faible sur des plants de pomelos (*Citrus grandis*). Cette différence d'agressivité ne se retrouvant pas sur les autres espèces de *Citrus*.

Pour trouver le gène responsable de cette différence de pouvoir pathogène, des mutations par insertion aléatoire du transposon Tn5 ont été effectuées dans le génome d'une souche de *Xcc* d'agressivité faible. Lors de l'analyse des phénotypes de ces mutants, l'un d'entre eux a montré une restauration de l'agressivité.



**Figure 4 : diversité des TALEs par AFLP**

A : Southern-blot obtenu par Al-Saadi *et al* en 2007, montrant la présence d'un TALE de 17.5 répétitions pour toutes les souches de *Xcc*

B : Southern-blot obtenu lors de l'étude préliminaire par A. Escalon avec à gauche des souches de pathotype A et à droite de pathotype A\*. On peut voir que les souches A possèdent au moins 3 tailles de TALEs alors que les souches A\* en possèdent au maximum 2. On peut constater également qu'il est possible de tirer un trait entre la bande correspondant à un TALE de 17.5 répétitions des souches A et celle correspondant à un TALE de 18.5 répétitions chez les souches A\* montrant bien ainsi une différence significative de la taille de ces homologues.



L'analyse de ce mutant a montré que le transposon Tn5 était inséré dans un TALE: le gène *hssB3.0*. L'analyse de la séquence de *hssB3.0* a montré que ce gène est une chimère, due à la recombinaison entre 2 TALEs présents chez cette souche.

Shiotani et son équipe ont donc démontré qu'un TALE pouvait être impliqué dans la restriction de la gamme d'hôtes, et que des événements de recombinaison entre les TALEs pouvaient se produire, formant ainsi de nouveaux TALEs fonctionnels potentiellement impliqués dans des modifications de la gamme d'hôtes.

### 1.4.3 Diversité des TALEs chez *Xcc*

La diversité des TALEs chez *Xcc* a été étudiée par Southern-blot (Al-Saadi *et al.* 2007 ; Lee, S *et al* 2008) et différents profils sont caractérisés par des contenus variables en nombres et en tailles (caractérisées par le nombre de répétitions supposées de la région centrale) de TALEs. A ce stade, aucune conclusion ne peut être tirée sur le nombre d'homologues présents dans une souche, sachant que chaque bande, ou taille de TALE, peut potentiellement correspondre à plusieurs homologues de séquences différentes.

L'analyse des profils de taille de TALEs (figure 4) a mis en évidence que les souches de pathotype A\* et A<sup>w</sup> ne possèdent que une à deux tailles de TALE différentes, alors que les souches de pathotype A en possèdent au minimum trois. Il est donc intéressant de comprendre si la différence de contenu en TALEs entre les pathotypes est responsable de la différence de gamme d'hôte.

Lors des études préliminaires de ce projet (A. Escalon), l'analyse des contenus en TALEs chez *Xcc* a été effectuée par Southern blot sur une collection de 56 souches représentatives de la diversité génétique de *Xcc*. Les résultats obtenus (confirmés par PCR) ont confirmé les observations faites par Al-Saadi et Lee, les souches du pathotype A étant caractérisées par un plus grand nombre de tailles de TALEs que les souches du pathotype A\*.

Dans ces travaux préliminaires, la migration d'ADN sur gel d'agarose a été faite de façon plus fine (temps de migration plus long permettant une meilleure séparation des bandes de taille différente), ce qui a mis en évidence que contrairement aux résultats obtenus par Al-Saadi *et al* en 2007, les souches de pathotypes A\* ne possèdent pas de TALE de 17.5 répétitions, mais une taille de TALEs de 18.5 répétitions (figure 4). Il est donc nécessaire de vérifier que ces 2 homologues jouent un même rôle dans la formation du chancre, et s'il existe des différences notables du point de vue de la séquence de ces homologues (RVDs).



## 1.5 Projet et enjeux

Les phénomènes d'émergence de nouvelles maladies peuvent être le résultat de plusieurs facteurs: introduction d'un pathogène dans une nouvelle zone géographique, changements environnementaux favorisant la propagation du pathogène, ou encore ces émergences peuvent être dues à une modification de la gamme d'hôtes d'un pathogène via des mécanismes évolutifs permettant de surmonter les systèmes de défense de la plante. Les mécanismes évolutifs sous-jacents aux modifications de gamme d'hôtes d'un pathogène donné sont encore mal connus. Les comprendre permettrait de mieux appréhender les phénomènes d'émergence de maladies à l'origine de pertes considérables pour les agriculteurs.

Le chancre asiatique des agrumes est une maladie émergente qui affecte l'ensemble des espèces d'agrumes commercialisées dans la plupart des pays agrumicole. L'agent pathogène responsable de cette maladie, *Xanthomonas citri* pv. *citri*, est caractérisé par une subdivision en 2 pathotypes, l'un (pathotype A) à large gamme d'hôte et l'autre (pathotype A\*) à gamme d'hôte restreinte, ce qui en fait un modèle intéressant dans la compréhension des mécanismes de spécialisation de la gamme d'hôtes. Le pouvoir pathogène de *Xcc* dépend, comme la plupart des *Xanthomonas*, d'un arsenal de protéines injectées dans la cellule végétale, les effecteurs de type III.

L'implication dans la détermination de la gamme d'hôtes d'un certain type d'effecteurs, les TALEs (Transcription Activator Like Effectors) a été récemment montrée chez d'autres *Xanthomonas*, notamment chez *Xanthomonas oryzae* pv. *oryzae*, pathogène du riz. Les TALEs sont caractérisés par une structure particulière en régions répétées, ce qui leur confère un potentiel évolutif fort, et supporte l'hypothèse d'un rôle dans l'évolution du pouvoir pathogène et des phénomènes d'adaptation à l'hôte. Chez *Xcc* l'effecteur *pthA4*, TALE comptant 17.5 répétitions, a un rôle primordial dans l'agressivité (Al-Saadi *et al.* 2007). Par ailleurs, des travaux réalisés sur un ensemble de souches montrant un phénotype atypique caractérisé par une agressivité réduite sur pamplemoussier Tachibana ont mis en évidence que cette modification de la gamme d'hôtes était due au TALE *hssB3.0* retrouvé chez toutes ces souches (Shiotani *et al.* 2007). Ces résultats indiquent donc que les TALEs peuvent être impliqués dans la modification de la gamme d'hôtes chez *Xcc*. Les souches appartenant aux différents pathotypes de *Xcc* sont caractérisées par une grande diversité de contenu en TALEs (Al-Saadi *et al.* 2007; Lee, S *et al.* 2008; A. Escalon études préliminaires). Ainsi, le rôle de



la plupart de ces TALEs dans la détermination de la gamme d'hôtes reste à élucider. Le fonctionnement de ces protéines comme activateurs de transcription est désormais connu. Un code de reconnaissance des résidus variables de chaque répétition (RVDs correspondant aux 12e et 13e acides aminés de chaque répétition) pour les bases du promoteur végétal a été établie par Boch *et al.* 2009; Bogdanove *et al.* 2010 et Streubel *et al.* 2012. Il est désormais possible de déterminer la séquence ciblée par un TALE donné dans la région promotrice du gène à activer chez l'hôte. Le séquençage des TALEs permet donc de caractériser les boxes théoriques ciblées par les différents homologues d'une souche et, *in fine*, de déterminer si les TALEs d'un pathotype donné participent à la détermination de la gamme d'hôtes, en ciblant spécifiquement certains gènes qui favorisent la colonisation d'un hôte. Il a donc été choisi de travailler sur une collection représentative de la diversité de *Xcc*, et de séquencer l'ensemble des TALEs de ces souches pour accéder aux boxes théoriques ciblées par les différents homologues. La comparaison des boxes ciblées permettra dans un premier temps de déterminer si certains motifs sont spécifiques d'un pathotype donné, c'est à dire si les TALEs contenus dans des souches appartenant à un même pathotype sont susceptibles d'activer les mêmes gènes, ou bien si au contraire on retrouve des motifs communs entre des souches ayant des gammes d'hôtes différentes. Dans un deuxième temps, grâce à une collaboration avec l'équipe de Boris Szurek (IRD de Montpellier), les séquences de RVDs caractérisant chacun des TALEs seront traitées par un algorithme de prédiction *in silico* visant à déterminer les cibles les plus probables activées chez les agrumes (mandarine et orange). Cette approche permet de déterminer si 2 TALEs ciblant des boxes différentes activent un même gène ou non. Ainsi, l'analyse d'un grand nombre de TALEs provenant de souches de profils de taille de TALEs différents, représentant la diversité de *Xcc* permettra de mieux appréhender le rôle de ces TALEs dans la détermination de la gamme d'hôtes. En fonction des résultats de séquençage, un travail de mutagénèse par délétion sera effectué afin de déterminer le rôle des différents homologues et appuyer les hypothèses formulées lors de l'analyse des séquences.

Souche	numéro	Origine Géographique	hôte d'isolation	Année d'isolation	Pathotype	dérivé Rif
<i>Xanthomonas citri</i> pv. <i>citri</i>	C40	Réunion	<i>Citrus sinensis</i>	1988	A	
<i>Xanthomonas citri</i> pv. <i>citri</i>	CFBP 2900	Japon	<i>Citrus sp</i>	<1982	A	
<i>Xanthomonas citri</i> pv. <i>citri</i>	CFBP 2911	Pakistan	<i>Citrus sp.</i>	1984	A*	CFBP 2911 <sup>R</sup>
<i>Xanthomonas citri</i> pv. <i>citri</i>	IAPAR 306	Brésil	<i>C. sinensis</i>	1997	A	
<i>Xanthomonas citri</i> pv. <i>citri</i>	JF90-2	Oman	<i>C.aurantifolia</i>	1986	A*	JF90-2 <sup>R</sup>
<i>Xanthomonas citri</i> pv. <i>citri</i>	JF90-3	Oman	<i>C.aurantifolia</i>	1986	A*	
<i>Xanthomonas citri</i> pv. <i>citri</i>	JF90-8	Oman	<i>C.aurantifolia</i>	1986	A <sup>w</sup>	JF90-8 <sup>R</sup>
<i>Xanthomonas citri</i> pv. <i>citri</i>	JJ10-1	Rodrigues	<i>C.aurantifolia</i>	1985	A	
<i>Xanthomonas citri</i> pv. <i>citri</i>	JJ238-24	Thaïlande	<i>C.aurantifolia</i>	1989	A*	
<i>Xanthomonas citri</i> pv. <i>citri</i>	JK143-9	Thaïlande	<i>C. reticulata</i>	NA	A*	
<i>Xanthomonas citri</i> pv. <i>citri</i>	JK48	Arabie Saoudite	<i>C.aurantifolia</i>	1988	A*	
<i>Xanthomonas citri</i> pv. <i>citri</i>	JS555	Iran	<i>C.paradisi</i>	1997	A*	
<i>Xanthomonas citri</i> pv. <i>citri</i>	JS581	Iran	<i>C. limetta</i>	1997	A*	
<i>Xanthomonas citri</i> pv. <i>citri</i>	JS584	Iran	<i>Citrus sp.</i>	1997	A*	JS584 <sup>R</sup>
<i>Xanthomonas citri</i> pv. <i>citri</i>	LE20-1	Ethiopie	<i>C.aurantifolia</i>	2008	A*	LE20-1 <sup>R</sup>
<i>Xanthomonas citri</i> pv. <i>citri</i>	LE3-1	Ethiopie	<i>C.aurantifolia</i>	2008	A*	
<i>Xanthomonas citri</i> pv. <i>citri</i>	LG100	Inde	<i>Citrus sp.</i>	2006	A <sup>w</sup>	
<i>Xanthomonas citri</i> pv. <i>citri</i>	LG102	Bangladesh	<i>Citrus sp.</i>	2006	A	
<i>Xanthomonas citri</i> pv. <i>citri</i>	LG115	Inde	<i>Citrus sp.</i>	2007	A <sup>w</sup>	
<i>Xanthomonas citri</i> pv. <i>citri</i>	LH1-3	Pakistan	<i>Citrus sp.</i>	2010	A	
<i>Xanthomonas citri</i> pv. <i>citri</i>	LMG 9322	Floride	<i>C.aurantifolia</i>	1986	A	
<i>Xanthomonas citri</i> pv. <i>citri</i>	NCPPB 3562	Inde	<i>C. lemon</i>	1988	A	
<i>Xanthomonas citri</i> pv. <i>citri</i>	NCPPB 3608	Inde	<i>C.aurantifolia</i>	1988	A <sup>w</sup>	
<i>Xanthomonas citri</i> pv. <i>citri</i>	NCPPB 3610	Inde	<i>Poncirus trifoliata</i>	1988	A	
<i>Xanthomonas citri</i> pv. <i>citri</i>	X2003-3218	Floride	<i>Citrus sp.</i>	2003	A <sup>w</sup>	

**Figure 5 : Souches de *Xanthomonas citri* pv. *citri* utilisées pour cette étude**

Souches de Xcc utilisées pour cette étude provenant de notre laboratoire, du Dr. T. Schubert, et de collections internationales : CFBP: Collection Française de Bactéries Phytopathogènes (INRA Angers, France); IAPAR: Instituto Agronômico do Paraná, Londrina PR, Brazil; BCCM/LMG: Belgian Coordinated Collections of Micro-organisms, University of Ghent, Belgium; NCPPB: National Collection of Plant Pathogenic Bacteria, York, UK.

## 2 Matériel et méthodes

### 2.1 Matériel

#### 2.1.1 Agrumes

Les plants d'oranger (*C.sinensis*) et de limettier Mexicain (*C.aurantifolia*) sont cultivés sur terreau et placés en chambre climatique avec une photopériode de 12h de lumière à  $30^{\circ}\text{C} \pm 1^{\circ}\text{C}$  et 12h d'obscurité à  $26^{\circ}\text{C} \pm 1^{\circ}\text{C}$  et avec une humidité relative de  $95\% \pm 5\%$ . Les jeunes feuilles de ces plants ont été sélectionnées pour l'inoculation par infiltration.

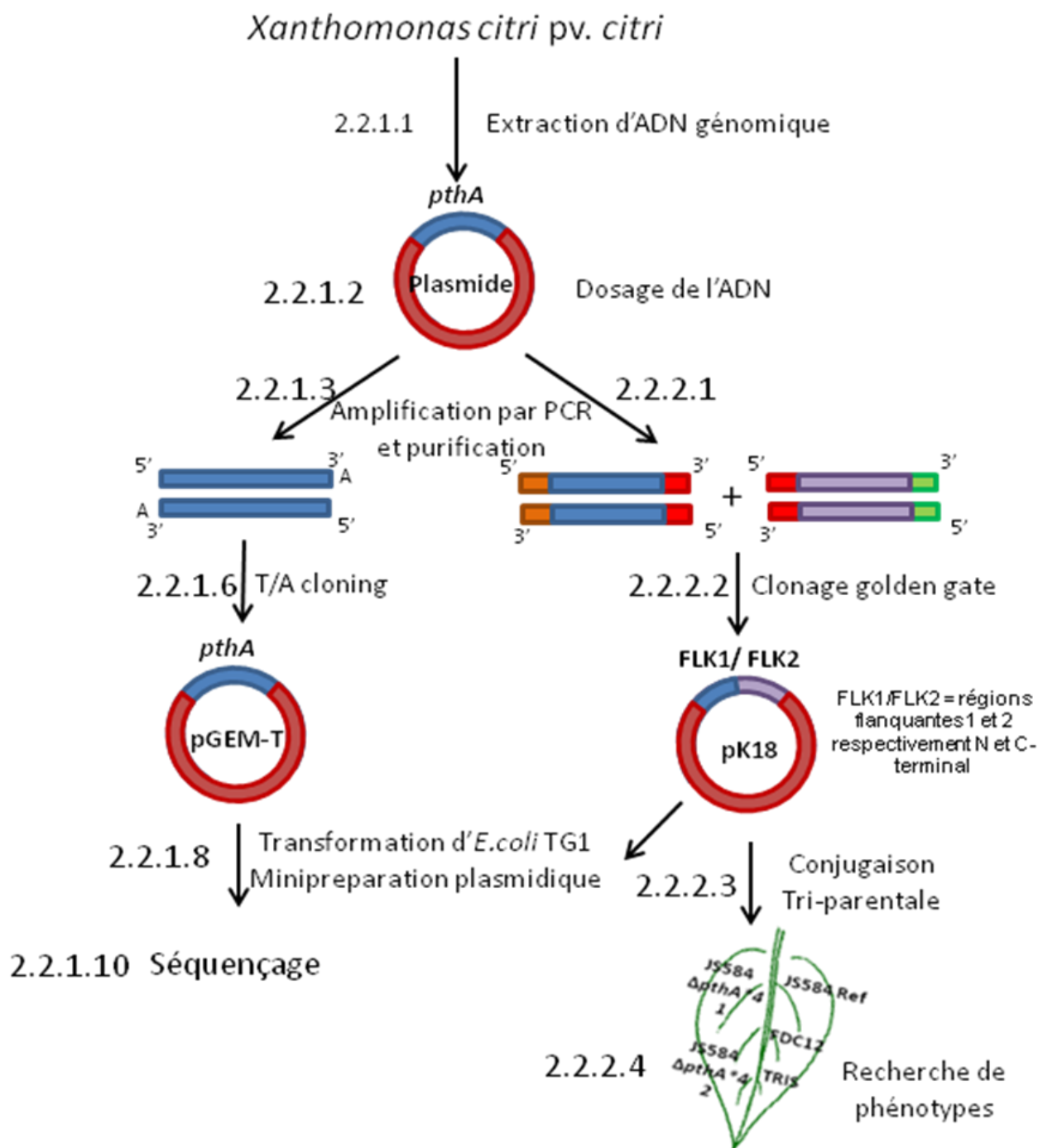
#### 2.1.2 Souches bactériennes et vecteurs

La souche bactérienne et les vecteurs utilisés pour les transformations et la multiplication des vecteurs sont les suivants :

- *Escherichia coli* TG1 F' [*traD36 proAB<sup>+</sup> lacI<sup>q</sup> lacZΔM15*]*supE thi-1 Δ(lac-proAB) Δ(mcrB-hsdSM)5, (r<sub>K</sub><sup>-</sup>m<sub>K</sub>)*.
- Plasmides :
  - pΔ13 plasmide pK18 compatible avec le clonage de type GoldenGate (L. D. Noël, non publié)
  - pRK2073 plasmide helper pour la conjugaison triparentale, ColE1 replicon, TraRK+, Mob+, Spec<sup>r</sup>
  - pGEM®-T easy plasmide utilisé pour le clonage des TALEs, Amp<sup>R</sup>, *lacZ*, promoteurs SP6 et T7, sites de reconnaissance des amorces universelles pUC/M13.

Les souches incluses dans cette étude ont été isolées de lésions chancreuses sur diverses variétés d'agrumes et collectées dans 13 pays différents. Elles ont été sélectionnées à partir de travaux précédents (Bui Ngoc *et al.* 2009) afin de maximiser la diversité en terme d'hôtes, d'année d'isolation, de pays d'origine ainsi que la diversité en terme de profil de TALEs. Elles sont listées dans le tableau figure 5.

Les dérivés rifampicine des souches destinées à la mutagénèse dirigée ont été sélectionnés en isolant les souches sauvages sur le milieu LPGA supplémenté à raison de 1000 µg/mL de rifampicine (cf figure 5). Pour chacune des souches, deux isolats résistants à la rifampicine ont été inoculés sur les agrumes pour vérifier leur pouvoir pathogène et ils ont été sélectionnés pour les expériences suivantes.



**Figure 6 : schéma général des grandes étapes des stratégies**

La colonne de gauche montre les étapes majeures de l'amplification et le clonage des TALEs expliquées dans le matériel et méthodes (paragraphe 2.2.1), jusqu'au séquençage pour l'analyse exploratoire des séquences. La colonne de droite montre les grandes étapes de la mutagenèse par délétion chez *Xcc* de pathotype A\*, jusqu'à l'inoculation sur plante des mutants pour la recherche de phénotypes (paragraphe 2.2.2).



## 2.2 Méthodes

### 2.2.1 Amplification et clonage des TALEs

La figure 6 résume les principales étapes de l'amplification au séquençage.

#### 2.2.1.1 Extraction d'ADN génomique

L'extraction d'ADN de *Xanthomonas* est faite à l'aide du kit Wizard® Genomic DNA Purification Kit (Promega) selon les recommandations du fournisseur, à partir d'une culture sur la nuit, sur milieu solide simple (LBA : extrait de levures 5g/L, pastone 10g/L, NaCl 10g/L, agar 16g/L et pH 7.5) ne contenant pas de sucre pour limiter la formation de polysaccharide.

#### 2.2.1.2 Dosage de l'ADN

La concentration de l'ADN est estimée par mesure spectrophotométrique au NanoDrop (ND-8000) à 260 nm, 1 unité d'absorbance correspond à une concentration de 50 µg/ml pour l'ADN double brin. Le degré de purification est estimé par les rapports DO260/DO280 dont la valeur doit être comprise entre 1.8 et 2 et DO260/DO230 dont la valeur doit être entre 1.8 et 2.2.

#### 2.2.1.3 PCR : Polymerase Chain Reaction

A partir de 20 ng d'ADN (amplification avant clonage) ou d'une colonie bactérienne reprise dans 200µL d'eau HPLC, Sigma Aldrich (vérification du clonage), l'amplification des fragments d'ADN est réalisée à l'aide d'amorces sens et anti-sens spécifiques à 0.3µM final. Les amorces utilisées pour l'amplification avant clonage (AE88-89) sont dessinées de manière à cibler la région centrale répétée des TALEs. La vérification des clones est faite à l'aide des amorces universelles M13 (figure 7). Le mélange réactionnel de la PCR (50 µL pour l'amplification avant clonage et 20µL pour la recherche de clones) contient: du tampon Green Flexi GoTaq® (1X), du MgCl<sub>2</sub> à 1.5mM, les 4 dNTPs à 0.2mM de chaque, la GoTaq® DNA polymérase (1.6U) et pour l'amplification avant clonage de la bétaine à 10 mM, destinée à faciliter l'amplification des séquences riches en GC. Toutes les amorces ont été dessinées à l'aide du logiciel primer3plus en ciblant une température de fusion (T<sub>m</sub>) de 60°C, l'étape d'hybridation est donc effectuée à 55°C et modifiée de plus ou moins 3°C en cas d'amplification non spécifique ou d'absence d'amplification due à une stringence trop importante.

amorce	sequence	oligo-spécifique	sites BsaI
AE88	GAGTTGAGAGGTCCACCGTTAC		
AE89	GAACCAATCGCCGTATTCCC		
AE60	AGCGGATAACAATTTACACAGG		
AE66	GTAAAACGACGGCCAGT		
SP6	ATTTAGGTGACACTATAG		
T7P	TAATACGACTCACTATAGGG		
1F	tttggcttcAAGGTCGGGCTTACGTTGCGACAC	CGGGCTTACGTTGCGACAC	tttggcttcAAGGT
2R	tttggcttcACATAACAGCGAGTTAGCTCACTTTTGGC	ACAGCGAGTTAGCTCACTTTTGGC	tttggcttcACATA
2F	tttggcttcATATGACCTGAGCGTCGGCAGGGAT	ACCTGAGCGTCGGCAGGGAT	tttggcttcATATG
3R	tttggcttcACTTGCGAAGATGCGGATGGGGCCG	CGAAGATGCGGATGGGGCCG	tttggcttcACTTG
E	TCCAGCGATGGGGCGCTTTC		
V1	GAAGTCAAGGGCGTAGAC		
V3	CAGAATTGTTCCGCACTCTG		
MG	GTTCTCGGCGGGCAGCTTCA		

### Figure 7 : Amorces utilisées

Les amorces AE88 et AE89 permettent l'amplification des fragments de TALEs à séquencer. Elles sont positionnées dans les régions conservées des TALEs proches des répétitions zéro et ½.

Les amorces AE60 et AE66 correspondent aux amorces universelles M13 et sont utilisées pour l'identification des clones contenant un TALE.

Les amorces universelles SP6 et T7 sont utilisées pour le séquençage à partir du plasmide pGEM®-T Easy.

Les amorces 1F à 3R sont les amorces utilisées pour le clonage golden-gate.

Enfin les amorces E à MG permettent l'identification des mutants *Xcc*.

Le temps d'élongation est choisi en fonction de la taille du fragment à amplifier. Sachant que la vitesse d'amplification de la GoTaq® polymérase est de 1minute/kb et que les fragments amplifiés les plus long font environ 2.7 kb, le temps d'amplification est de 3 minutes. Le protocole thermique utilisé est donc : 1. Dénaturation initiale à 95°C 4min; 2. Dénaturation à 95°C 40s; 3. Hybridation à 55°C 40s; 4. Elongation à 72°C 3 min; 5. Elongation finale 72°C 5 min; 6. Conservation à 12°C ∞. Les étapes 2 à 4 sont répétées 25 fois pour l'amplification qui précède le clonage et 30 fois pour la recherche de clones contenant le TALE d'intérêt.

#### **2.2.1.4 Electrophorèse de l'ADN sur gel d'agarose**

L'analyse des amplicons est effectuée sur des gels d'agarose de 1 ou 1.5 % dans du TAE 1 X (TRIS-acétate 40mM, EDTA 1mM). L'échelle de référence utilisée (100 pb DNA Step ladder, Promega) permet d'identifier la taille des fragments d'ADN. La migration électrophorétique est réalisée à voltage constant (160V) dans du tampon TAE 1X. Le gel est incubé dans une solution de BET (2 µg.mL<sup>-1</sup>). Une photographie est réalisée avec l'appareil Transluminateur ETX-F26.M, Caméra CFW-1300 et le logiciel PerfectImage V7.6.

#### **2.2.1.5 Purification des amplicons**

Les amplicons identifiés et destinés au clonage sont découpés sur gel d'agarose en prenant soin de minimiser le temps d'exposition aux UVs. Pour chaque souche, on dispose donc d'un nombre de fragments qui correspond à son contenu en TALEs de tailles différentes. Les fragments découpés sont purifiés à l'aide du kit Promega Wizard® SV Gel and PCR Clean-Up System, par fixation des molécules d'ADN sur une colonne de silice.

#### **2.2.1.6 Clonage des TALEs dans pGEM®-T Easy**

Le clonage des fragments purifiés se fait avec le kit Promega pGEM®-T Easy Vector Systems. La pureté et la qualité d'insert est vérifiée sur gel d'agarose 1%, puis un volume de 3µL d'insert est ajouté à 50ng de vecteur pGEM®-T Easy, du tampon de ligation 1X et de la T4 DNA ligase à 3U dans un volume final de 10 µL selon les recommandations du fournisseur. Le mélange réactionnel est incubé à 4°C pendant 48h. Le mélange est fait de manière à contenir un maximum d'insert.

#### **2.2.1.7 Electroperméabilisation d'*E.coli* TG1**

A partir d'une culture d'*E.coli* TG1 sur la nuit en milieu LB liquide, 200 mL de milieu SOB liquide sontensemencés au 1/100<sup>èmes</sup>. La culture est suivie jusqu'à une DO<sub>600nm</sub> d'environ 0.6, (soit environ 3h à 37°C) puis mise dans la glace. Toutes les étapes suivantes



sont effectuées à froid, à 4°C. 1. Centrifugation 15 minutes à 6000 rpm. 2. Lavage par reprise dans 200 mL d'eau stérile. 3. Centrifugation 15 minutes à 6000 rpm. 4. Lavage avec 100 mL d'eau stérile. 5. Centrifugation 15 minutes à 6000 rpm. 6. Reprise du culot dans 10 mL de glycérol à 10% stérile. 7. Centrifugation 15 minutes à 4000 rpm. 8. Reprise du culot dans 0.8 mL de glycérol à 10% stérile. Les cellules électrocompétentes sont alors aliquotées par 50µL et congelées à - 80°C.

#### **2.2.1.8 Transformation de bactéries compétentes et préparation plasmidique**

Après dialyse, le mélange de ligation (5 µL) est mis en contact avec 50µL de bactéries électrocompétentes. La transformation est faite par choc électrique (2500V, 200 ohm, 25µF), les cellules sont rapidement reprises avec 1mL de milieu LB liquide, puis incubées 1h à 37°C. 100µL de la suspension et 100µL du culot sont ensuite étalés sur milieu LB à 100 µg/mL d'ampicilline, 80 µg/mL X-gal, 0.5 mM d'IPTG. Les boîtes de Petri sont ensuite incubées à 37°C sur la nuit. Le lendemain, huit colonies isolées sont repiquées sur une nouvelle boîte de milieu solide LB+ampicilline+X-gal+IPTG et testées par PCR sur colonies à l'aide des amorces M13 afin de vérifier la présence et la taille de l'insert. Quatre clones sont choisis pour la suite des manipulations: d'une part, repiquage pour conservation en microbilles à - 80°C, et d'autre part, préparations plasmidiques destinées au séquençage. Les clones, sont remis en culture dans 5 mL de LB liquide +ampicilline à 100µg/mL et les bouillons sont incubés sur la nuit à 37°C sous agitation. On effectue l'extraction plasmidique sur ces cultures liquides, à l'aide du kit Promega Wizard® Plus SV Minipreps DNA Purification System par lyse bactérienne et fixation des molécules d'ADN sur colonne de silice sous des conditions chaotropiques en sels. Après plusieurs lavages pour éliminer les contaminants, l'ADN plasmidique pur est élué dans 50µL d'eau HPLC.

#### **2.2.1.9 Vérification de la qualité des extractions plasmidiques**

Afin de contrôler la qualité des minipréparations, et de vérifier la taille de nos inserts, 2µL de chaque minipréparations sont digérés en 5' et en 3' de l'insert à l'aide d'EcoRI avec 20 U d'enzyme de restriction, 1X de tampon correspondant à l'enzyme, 10µg de BSA et de l'eau HPLC pour 10µL final. Le tout est incubé à 37°C pendant 3h, et visualisé après électrophorèse sur gel d'agarose à 1.5%.

En parallèle les minipréparations sont dosées au nanodrop de façon à contrôler leur pureté (cf 2.2.1.2). La cohérence entre la concentration et l'intensité est visualisée sur le gel.



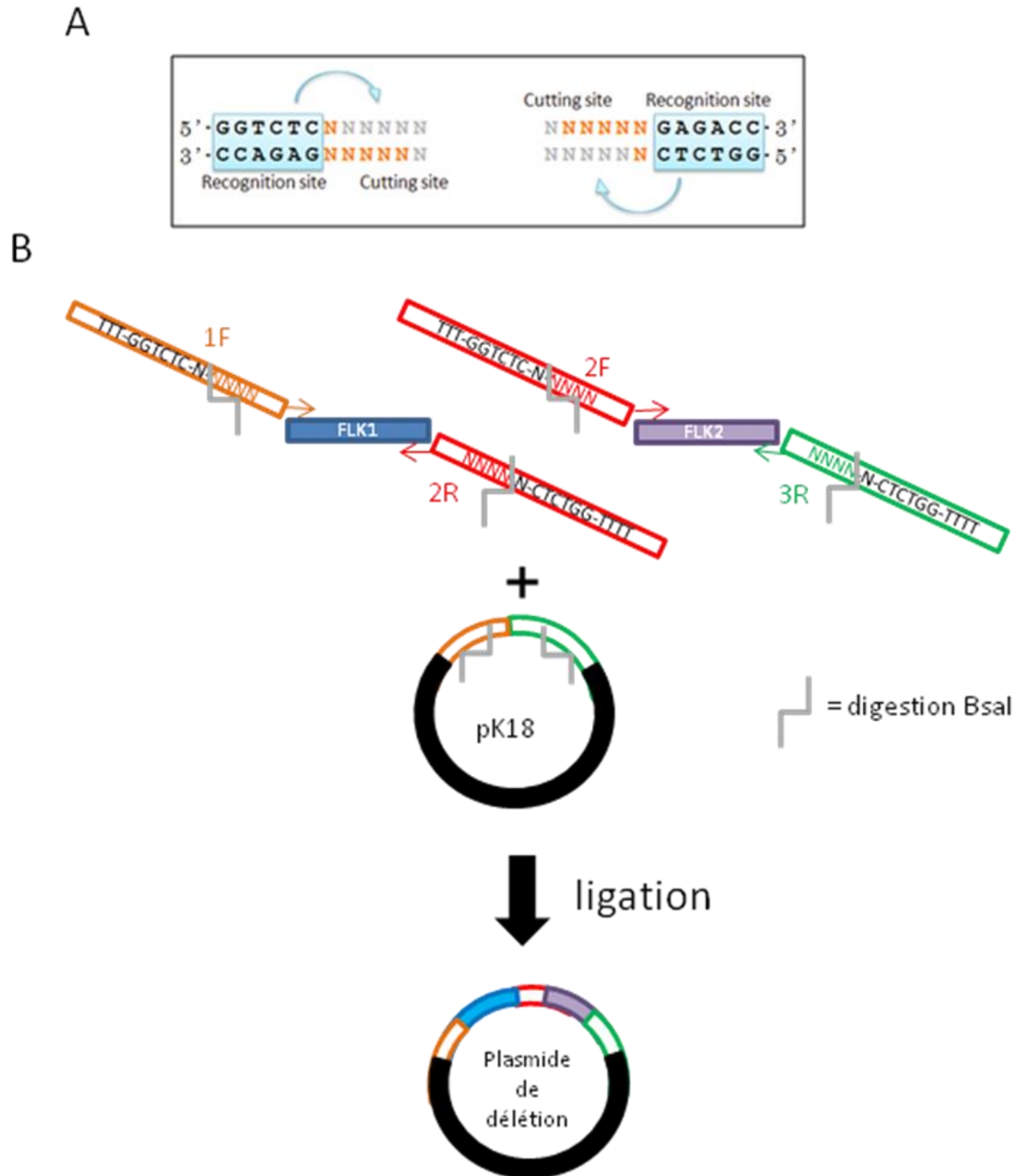
### 2.2.1.10 Séquençage

Le séquençage des TALEs est fait dans un premier temps en utilisant les amorces SP6 et T7 universelles (figure 7). Le séquençage est fait par l'entreprise Beckman Coulter Genomics, qui utilise un séquenceur Applied Biosystems ABI 3730xl pour le séquençage Sanger. Le résultat est analysé grâce au logiciel Geneious: qualité des chromatogrammes, assemblage des runs et lecture des RVDs (Repeat Variable Di-residus). Dans un second temps, des amorces spécifiques sont dessinées de manière à séquencer les répétitions centrales non couvertes par les premiers runs de séquençage.

## 2.2.2 Délétion du gène *pthA\*4*

### 2.2.2.1 Amplification des régions flanquantes et purification

A partir de l'ADN des souches à muter (100 ng), l'amplification des régions flanquantes du gène *pthA\*4* est réalisée à l'aide d'amorces sens et anti-sens spécifiques dessinées dans ces mêmes régions du gène *pthA4* de la souche de référence *Xcc 306* à 0.3µM final. Le mélange réactionnel de la PCR (50 µL) contient: du tampon HF (1X), les 4 dNTPs à 0.2mM de chaque et l'enzyme PHusion® High Fidelity DNA polymerase (Finnzymes). Le temps d'élongation est choisi en fonction de la taille du gène à amplifier, sachant que la vitesse d'amplification de l'enzyme PHusion® est de 15 à 30sec/kb. Le protocole thermique utilisé est : 1. Dénaturation initiale à 98°C 2min; 2. Dénaturation à 98°C 20s; 3. Hybridation à 58°C 30s; 4. Elongation à 72°C t min; 5. Elongation finale 72°C 4 min; 6. Conservation à 12°C ∞. Les étapes 2 à 4 sont répétées 35 fois. La taille et la concentration des amplicons sont vérifiées sur gel d'agarose. Les produits PCR découpés après électrophorèse sur gel d'agarose sont purifiés à l'aide du kit Promega Wizard®SV Gel and PCR Clean-Up System, par fixation des molécules d'ADN sur une colonne de silice.



**Figure 8 : Principe du clonage Golden-gate**

A : sites de reconnaissances et de coupure de l'enzyme de restriction de type IIS : BsaI, générant des extrémités 5' débordantes.

B : La PCR des régions flanquantes se fait à l'aide d'amorces à queue flottante, possédant le site de restriction de BsaI, une série de thymines en 5' pour stabiliser l'enzyme et un site en 3' générant des extrémités cohésives au moment de la ligation pour un clonage orienté.



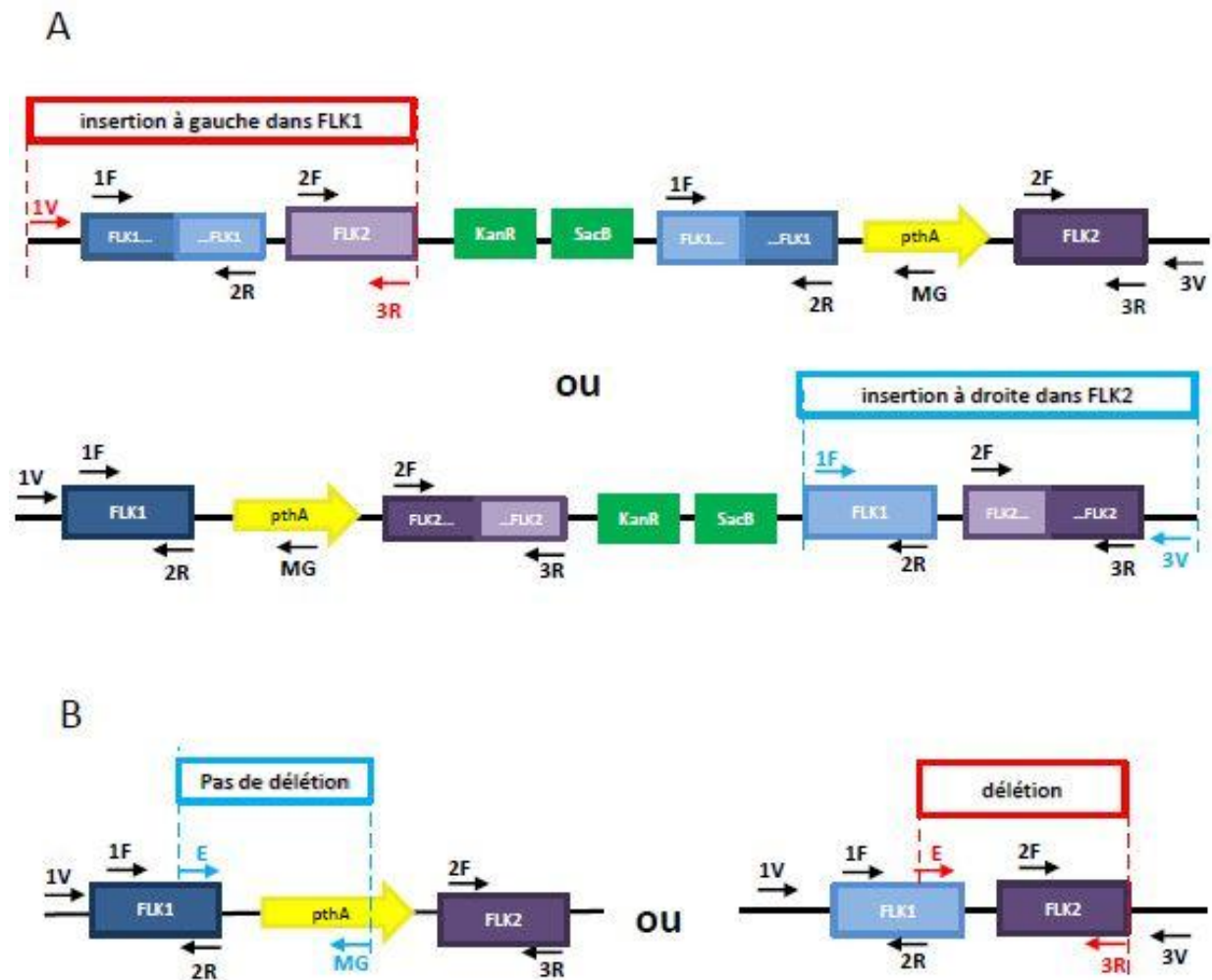
### 2.2.2.2 Clonage Golden-gate

Ce clonage orienté repose sur une série de digestion/ligation en chaîne, cela étant possible grâce à l'utilisation d'une enzyme de restriction de types IIS, c'est-à-dire qui coupe à l'extérieur de son site de reconnaissance. On utilise pour la digestion, l'enzyme BsaI à 30 U qui reconnaît la séquence palindromique GGTCTC en 5' et CCAGAG en 3' et qui coupe respectivement 1 et 5 bases après ce site de reconnaissance (figure 8A). On a donc après digestion la perte du site de restriction de BsaI d'où l'enchaînement possible des étapes de digestion et de ligation par la ligase T4 à 6 U, au cours desquelles le mélange réactionnel s'enrichit progressivement en vecteur+insert dépourvu du site de reconnaissance de l'enzyme. L'amplification des régions flanquantes 1 (FLK1) et 2 (FLK2) (cf 2.2.2.1) a été réalisée respectivement à l'aide des couples d'amorces 1F-2R et 2F-3R (figure 7) dessinées avec des queues flottantes contenant le site BsaI ainsi qu'un fragment spécifique de 4 pb permettant l'orientation de ce clonage au moment de la ligation: orientation des fragments FLK1 et FLK2 et ligation complémentaire avec les sites libérés par la digestion du plasmide pΔ13 (figure 8B).

Le mélange réactionnel de ce clonage (15 μL) contient : 1 μL de vecteur pΔ13 (minipréparation à partir *E.coli* TG1 contenant ce vecteur), 3μL d'amplicons (de chaque région flanquante) purifiés, du tampon de ligation T4 à 1X, de la T4 DNA ligase à 6 U, de la BSA au minimum facilitant la digestion et la BsaI à 30 U.

Le programme thermique utilisé est : 1. Digestion initiale par BsaI à 37°C 5min ; 2. Digestion par BsaI à 37°C 5 min ; 3. Ligation par la T4 ligase à 16°C 5 min ; 4. Inactivation de BsaI à 50°C 5 min ; 5. Inactivation de la T4 ligase à 80°C 10 min ; 6. Conservation à 20°C ∞. Les étapes 2 et 3 sont répétées 35 fois.

Les plasmides de délétions ainsi obtenus sont ensuite dialysés et électroporés dans *E.coli* TG1 (cf 2.2.1.8). Les bactéries transformées sont étalées sur milieu LB à 25 μg/mL de kanamycine antibiotique de sélection dont la résistance est apportée par le vecteur pΔ13. Le lendemain, huit colonies isolées sont repiquées sur une nouvelle boîte de milieu solide LB+kanamycine et testées par PCR sur colonies à l'aide des amorces universelles M13 afin de vérifier la présence et la taille de l'insert. Quatre clones sont choisis pour la suite des manipulations: d'une part, repiquage pour conservation en microbilles à -80°C, et d'autre part, conjugaison tri-parentale comme souche donneuse du plasmide de délétion.



**Figure 9 : Positionnement des amorces pour la recherche de mutant *Xcc* par délétion**

A : Recherche de l'insertion du plasmide de délétion dans l'une des deux régions flanquantes (FLK) par recombinaison homologue. Obtention d'une bande sur gel d'agarose après électrophorèse : avec le couple 1V-3R si le plasmide s'est inséré dans la région flanquante 1, et avec 1F-3V pour une recombinaison dans la région flanquante 2.

B : Recherche de mutants délétés du gène *pthA*\*4. Si le gène n'a pas été délété une bande sera visible sur gel après une amplification avec le couple E-MG. En revanche si la délétion s'est produite une bande sera visible avec le couple E-3R.

### 2.2.2.3 Conjugaison tri-parentale et sélection des mutants

A partir de culture de 24h, les souches nécessaires à la conjugaison tri-parentale sont mises en culture (bouillons): LB+kan pour les souches d'*E.coli* donneur, LB+Spec pour la souche *E.coli* Helper (pRK2073), MOKA (extrait de levure : 4g/L ; pastone : 8g/L ;  $K_2HPO_4$  : 2g/L ;  $MgSO_4 \cdot 7H_2O$  : 0.3g/L ; agar : 18g/L et pH 7.2) pour les souches de *Xanthomonas* à transformer. Les cultures sont lancées sur la nuit sous agitation à 37°C pour *E.coli* et 28°C pour *Xanthomonas*. Le lendemain, 1mL de chaque culture est centrifugé 5 minutes à 5000 rpm. Chaque culot est ensuite repris dans 600 µL de milieu liquide (LB pour *E.coli* et MOKA pour *Xcc*), puis 60 µL de chacune de ces suspensions sont mélangés et déposés sur un filtre (cellulose, pores de 40 µm de diamètre) dans une boîte de MOKA. Après évaporation du liquide sur le filtre les boîtes sont incubées à 28°C sur la nuit.

Pour sélectionner les transformants, la culture ayant poussé sur le filtre est resuspendue dans 1 mL d'eau milliQ stérile, puis centrifugée 5min à 5000 rpm. 800 µL de surnageant sont éliminés et le culot est repris dans les 200 µL restants, puis étalé sur milieu doublement sélectif MOKA à 100 µg/mL de Rifampicine et 50 µg/mL de Kanamycine. Les boîtes sont incubées à 28°C pendant 3 à 4 jours. 12 colonies sélectionnées par leur aspect (jaunes, brillantes, à bord lisse et semi-bombées) sont repiquées successivement sur différents milieux, de façon à vérifier leur phénotype: 1. LBA pour l'aspect et couleur de la colonie ; 2. MOKA+Rif sauvegarde pour les étapes suivantes et aspect de la colonie ; 3. MOKA+ Rif+ Kan où seuls les *Xanthomonas* ayant inséré le plasmide de délétion poussent grâce au gène *Kan<sup>R</sup>* de pΔ13; 4. MOKA+ Rif +Kan+ Sucrose où les *Xcc* possédant le plasmide de délétion ne peuvent pas pousser à cause de la cassette SacB du vecteur. Toutes les boîtes sont incubées à 28°C sur la nuit.

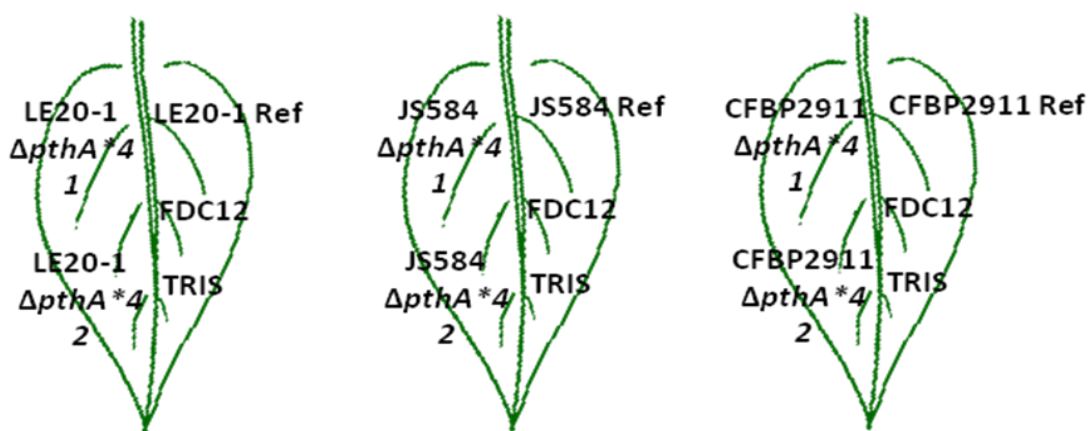
A partir de la sauvegarde sur MOKA+Rif+Kan, les colonies ayant donné le bon phénotype sont testées en PCR avec 2 couples d'amorces : V1-3R pour voir l'insertion dans FLK1 et 1F-V3 pour l'insertion dans FLK2 (figures 7 et 9A). Les colonies pour lesquelles l'insertion du plasmide de délétion est confirmée sont ensuite isolées sur milieux MOKA +Rif + sucrose de façon à induire une seconde recombinaison homologue pour expulser le gène *pthA*\*4 à déléter.

Enfin, la délétion est vérifiée sur 12 colonies isolées avec les couples d'amorces E-3R permettant de voir la délétion et E-MG montrant la présence du gène (figures 7 et 9B). 3 mutants par souche sauvage A\* ainsi sélectionnés sont ensuite conservés sur billes à -80°C dont 2 sont testés par inoculation sur plante pour l'étude du phénotype correspondant.



#### 2.2.2.4 Caractérisation phénotypique des mutants

Les jeunes feuilles de limettiers et d'oranger ont été inoculées, à  $10^5$  CFU/mL, par infiltration de la face inférieure. Les suspensions sont préparées à partir de cultures fraîches dans du TRIS sachant qu'une  $DO_{\lambda=600nm}$  de 0.06 correspond à  $10^8$  CFU/mL. Pour chaque souche sauvage A\*, 2 mutants  $\Delta PthA^*4$  sont inoculés. Les 2 mutants, la souche sauvage, la souche du pathotype A IAPAR 306 (témoin positif) et une solution de TRIS (témoin négatif) sont inoculés sur une même feuille. 5 répétitions (A à E) réparties sur 4 plants différents sont effectuées, en respectant le plan d'inoculations suivant :



L'évolution des symptômes est suivie tous les 3 jours jusqu'à J+25 après l'inoculation.

#### 2.2.2.5 Extraction des bactéries sur symptômes et vérification des génotypes par PCR.

A J+21 après inoculation, L'ensemble des feuilles de la répétition "A" est détachée pour vérification des génotypes. Chaque zone infiltrée est découpée, puis dilacérée dans 1 mL TRIS. 50μL du macérât sont isolés sur LPGA (extrait de levure : 7g/L ; glucose : 7g/L ; pastone : 7g/L ; agar : 18 g/L ; Tilt : 80μL/L et pH 7.2, ce milieu permettant une meilleure croissance des *Xanthomonas* que le milieu MOKA) à 100 μg/mL de Rifampicine. Les boîtes sont ensuite incubées 3 jours à 28°C.

Huit colonies isolées sont analysées par PCR avec les couples d'amorce E-3R, E-MG et AE60-66. Cela permettant de contrôler l'absence de *pthA^\*4* chez les mutants et sa présence chez la souche non mutée correspondante (Rif).



## 3 Résultats et discussion

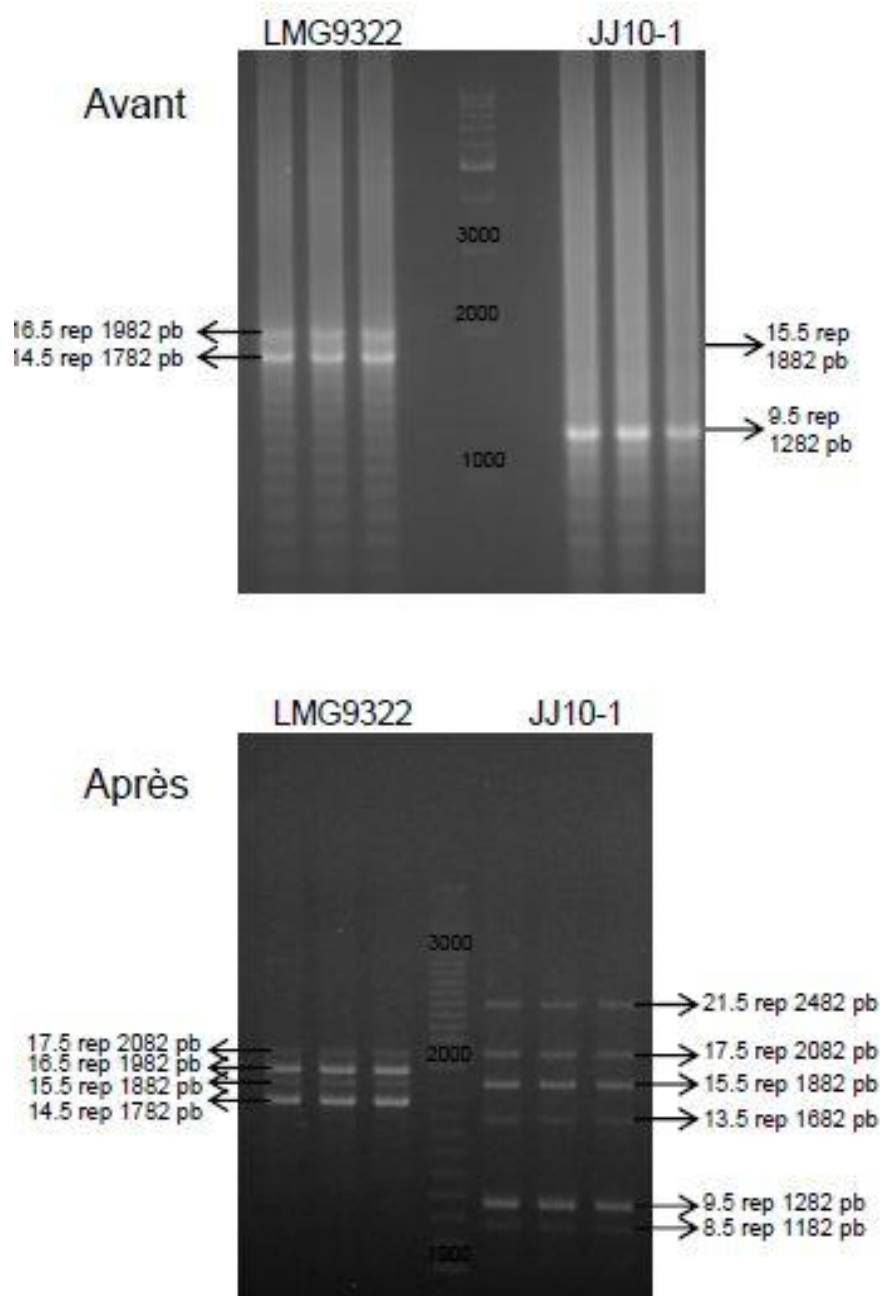
### 3.1 Caractérisation des TALEs chez *Xanthomonas citri* pv. *citri*

De nombreux travaux ont montré l'implication des TALEs dans l'agressivité de *Xcc* (Swarup 1991, Swarup 1992, Al-Saadi *et al* 2007), et plus récemment leur implication dans la limitation de la gamme d'hôtes (Shiotani *et al* 2007). L'étude de la diversité des TALEs chez *Xcc* a mis en évidence différents profils caractérisés par des contenus variables, en nombre et en tailles, de TALEs en fonction du pathotype des souches (Al-Saadi *et al* 2007, Seungdon Lee *et al.* 2008, études préliminaires A.Escalon par PCR et Southern blot). Par ailleurs, la découverte du code permettant d'associer à un TALE donné une séquence promotrice ou boîte cible spécifique (Boch *et al* 2009, Bogdanove *et al* 2009), implique qu'au delà de la diversité liée au nombre de répétitions qui caractérisent les TALEs, il est important de déterminer la séquence spécifique ciblée par chacun d'eux. Une étude approfondie de la diversité des TALEs de *Xcc* paraît intéressante afin de déterminer leur rôle dans la spécialisation de la gamme d'hôte entre les différents pathotypes de *Xcc*. Le but de cette étude est donc d'identifier les séquences promotrices fixées par chaque TALE, afin de pouvoir les comparer dans un premier temps, et tenter de déterminer quels sont les gènes activés par les différents homologues.

Une sous collection de 20 souches a été choisie afin de représenter au mieux la diversité des contenus en TALEs, la diversité génétique de *Xcc*, et les origines géographiques des souches. La stratégie suivie étant de séquencer l'ensemble des TALEs de chacune des souches de cette sous collection.

#### 3.1.1 Mise au point des différentes étapes

Plusieurs mises au point des protocoles ont été nécessaires au cours des différentes manipulations.



**Figure 10 : Mise au point de l'amplification des TALEs par PCR**

Gels d'agarose 1%, des souches LMG9322 et JJ10-1, avant et après la mise au point de la PCR pour l'amplification des TALEs. Dans les 2 cas les PCR sont faites en 3 exemplaires pour augmenter la quantité d'inserts pour le clonage. Après la mise au point toutes les bandes attendues sont visibles, de bonne intensité et non dissimulées dans le smear.



### 3.1.1.1 Amplification par PCR

L'amplification des TALEs par PCR est délicate, du fait de leur composition en régions répétées. Le protocole d'amplification a été optimisé de manière à palier aux phénomènes inhérents à l'amplification de régions répétées, présentes en multicopies dans une même souche: 1. les séquences répétées peuvent s'apparier avec un décalage de une à plusieurs répétitions lors des différents cycles de PCR, donnant ainsi après électrophorèse sur gel d'agarose, un smear caractéristique de l'amplification des TALEs; 2. les différents TALEs d'une souche donnée ne sont pas tous amplifiés avec le même rendement et la concentration d'ADN initiale peut influencer sur le rendement d'amplification pour certains homologues.

La mise au point a par conséquent visé à déterminer pour chaque souche la concentration initiale optimale d'ADN à utiliser pour les PCR, de façon à obtenir un compromis entre une amplification suffisante de toutes les tailles de TALEs attendues, et une limitation du smear. Les concentrations initiales variant pour nos souches de 0.005 à 1 ng/ $\mu$ L.

Les premiers résultats de séquençage ont montré la présence de séquences parasites (double signal dans le chromatogramme) provenant probablement du smear lors du découpage des bandes, et/ou d'appariements inadéquats lors des cycles de PCR.

Une diminution du nombre de cycles, de 35 à 25, a été testée de façon à limiter les erreurs de la PCR. Les résultats ont montré une absence de smear, mais également une très faible intensité des bandes. Une nouvelle détermination de la concentration optimale a donc été effectuée; 5 ng/ $\mu$ L ayant donné de bons résultats pour la majorité des souches (résultat figure 10)

### 3.1.1.2 Purification des produits PCR et ligation dans pGEM®-T Easy

L'amplification par PCR des TALEs est faite de manière à avoir un volume total de 200 $\mu$ L d'amplicons pour chaque souche (multiplication du nombre de puits), ce qui permet d'augmenter la quantité d'insert obtenue après purification des fragments découpés sur gel. On dispose ainsi d'une concentration suffisante en insert pour la ligation. Des étapes ont été rajoutées à la fin du protocole du kit Promega. Une étape de centrifugation 5 minutes à 14 000 rpm, après le second lavage à l'éthanol ; suivie d'un temps de pause, de 20 minutes, à l'air libre colonne retournée. Ces étapes permettent une évaporation complète de l'éthanol.

Plusieurs temps de ligation dans le pGEM®-T Easy, ont été testés : 4h à 37°C, 4h à température ambiante, 24h et 48h à 4°C. Une ligation de 48h à 4°C (ce qui est conforme aux recommandations du fournisseur), donne le meilleur rapport de colonies blanches sur colonies



bleues, mais également une meilleure qualité de clonage (nature de l'insert confirmée par PCR sur colonies).

L'étape de la dialyse avant électroporation permet par ailleurs de limiter fortement les explosions lors du choc électrique, explosions dues à la présence de sels dans le tampon de ligation.

### **3.1.1.3 Minipréparations de plasmides pour séquençage.**

Les résultats préliminaires ont mis en évidence l'importance de la qualité et de la concentration des minipréparations pour obtenir des résultats de séquençage exploitables en termes de qualité et de longueur de runs.

La concentration minimum de 100 ng/μL requise pour le séquençage a été obtenue en lançant les cultures en bouillon LB à partir de repiquage frais du matin et en augmentant la concentration en ampicilline de 50μg/mL à 100μg/mL pour maintenir une meilleure pression de sélection. De plus, plusieurs volumes de culture utilisés pour la minipréparation ont été testés : 1.5, 2, 3 et 4 mL ; le meilleur rendement ayant été obtenu en centrifugeant 4 mL de la culture.

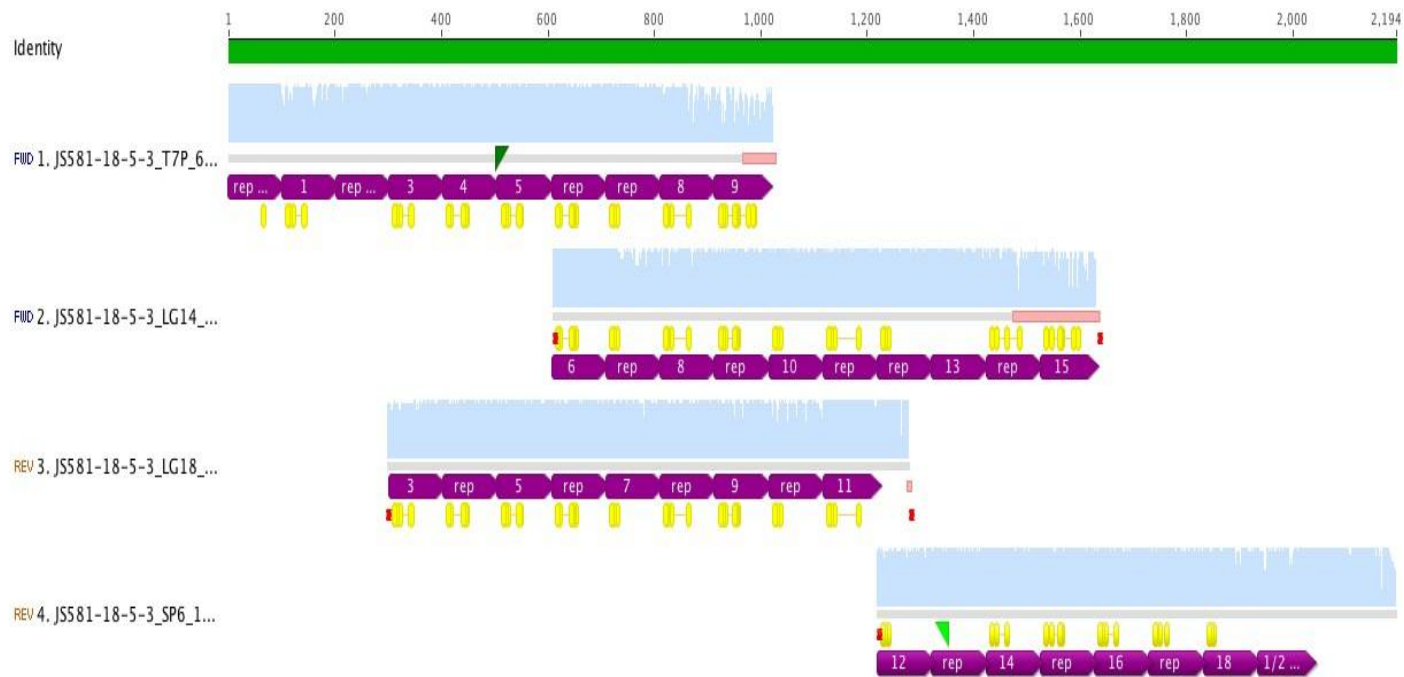
Enfin, les dernières étapes du protocole (annexe 1) ont été modifiées de la même manière que le protocole de purification des produits de PCR : évaporation complète de l'éthanol avec une centrifugation supplémentaire et une étape de séchage. Par ailleurs, l'étape d'incubation avec l'eau HPLC pour l'élution finale a été rallongée de 5 minutes avant la dernière centrifugation.

## **3.1.2 Résultats de séquençage**

Les chromatogrammes obtenus ont été analysés à l'aide du logiciel Geneious et comparés avec les TALEs de la souche 306 de référence entièrement séquencée.

### **3.1.2.1 Recherche d'homologues et relevé des RVDs**

Chaque chromatogramme est analysé de la manière suivante : recherche du motif de 99 à 102 pb caractéristique des TALEs à l'aide de l'outil de recherche de motifs, annotations de la répétition "zéro" ainsi que de la dernière demi-répétition. Pour chaque taille de TALEs de chaque souche, au moins 3 minipréparations ont été envoyées à séquencer, ce qui permet de vérifier par alignement s'il existe, pour un fragment donné, des homologues distincts du point



**Figure 11 : Séquençage des répétitions centrales**

Un des TALEs de 18.5 répétitions séquençé de la souche JS581, est utilisé ici en exemple pour montrer la nécessité d'un second run de séquençage pour obtenir les répétitions centrales.

Les premiers runs (1 et 4) ont permis de séquencer environ 1000 pb chacun, cela ne suffisant pas pour obtenir la séquence complète d'un TALE d'environ 2200pb. De façon à séquencer les répétitions manquantes (ici 10 et 11) des amorces (triangles verts) ont été dessinées dans les répétitions montrant un ou des SNPs spécifiques (ici 5 et 13). Les séquences 2 et 3 ont été respectivement obtenues à partir de ces amorces. Un assemblage des quatre chromatogrammes (en bleue) a pu être effectué, donnant ainsi la séquence complète de ce TALE.

de vue de leur composition en RVDs. Les résultats de ces alignements ont mis en évidence à plusieurs reprises la présence d'homologues différents (et donc de RVDs différents) de même taille (annexe 2).

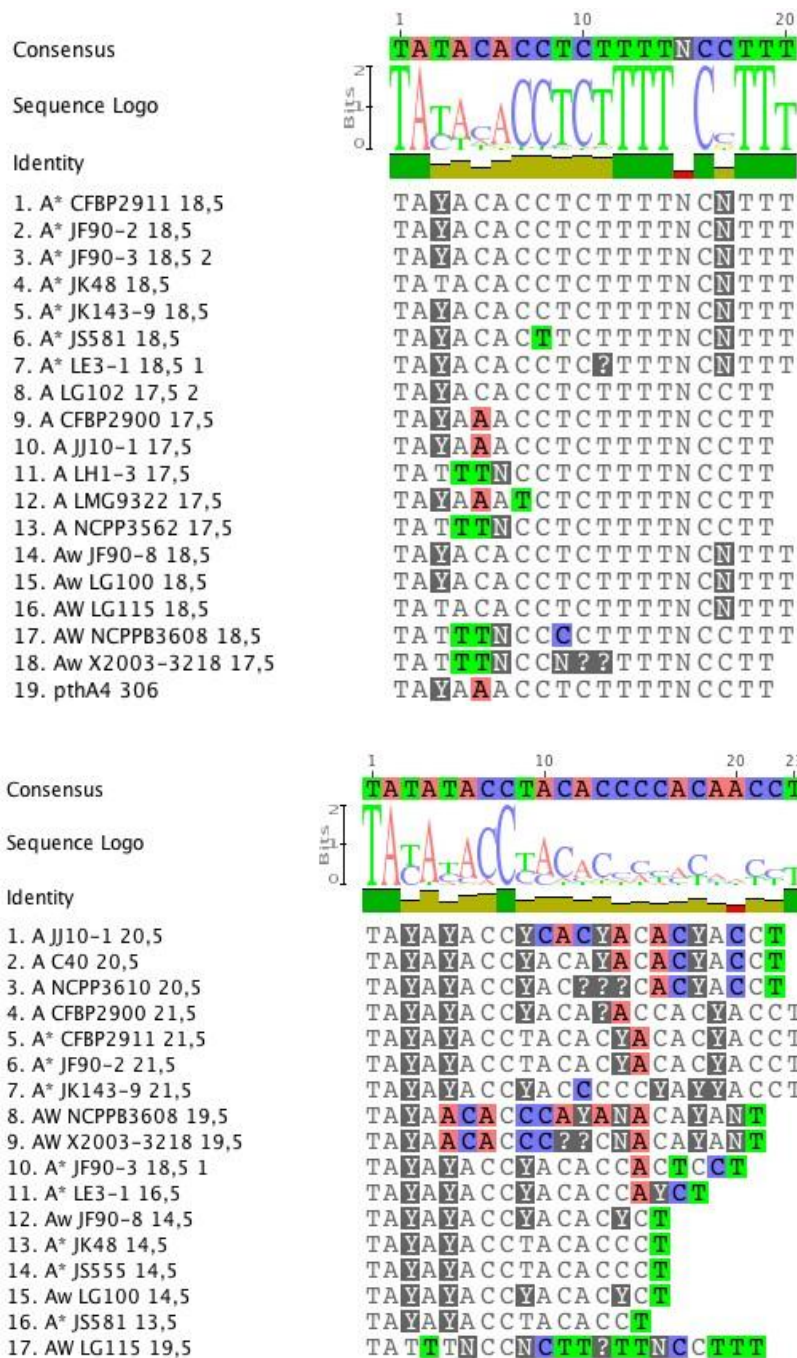
La présence de 2 homologues différents pour une même taille de TALE a été retrouvée chez trois souches du pathotype A : LG102, LMG9322 et CFBP2900 ; et ce pour une seule taille de TALE par souche, augmentant ainsi la diversité des TALEs de ce pathotype par rapport aux souches du pathotype A\*. Chez les souches du pathotype A\*, la présence de 2 homologues de taille similaire n'a été mise en évidence que pour la souche JF90-3 caractérisée par une seule taille de TALEs à 18.5 répétitions (toutes les autres souches A\* étant caractérisées par un profil à 2 bandes, soit 2 tailles de TALEs). Le séquençage a donc mis en évidence que toutes les souches A\* possèdent au moins 2 TALEs.

Les RVDs de chaque homologue séquencé sont répertoriés dans le tableau de l'annexe 1, avec pour chaque TALE le nombre de plasmides séquencés ayant donné le même résultat de séquençage.

On remarque que lorsque l'on a deux homologues de même taille, l'un d'entre eux est retrouvé à fréquence plus faible, cela pouvant être dû à un nombre de copie plus faible initialement dans la souche de départ.

### 3.1.2.2 Design d'amorces pour les répétitions centrales

Les runs de séquençage réalisés à partir des amorces T7 et SP6 ont en général donné des chromatogrammes analysables sur 800 à 1000 pb. Par conséquent, la plupart des TALEs de moins de 17.5 répétitions et donc les amplicons de longueur totale inférieure à 2100 pb, ont pu être entièrement séquencés dès le premier run. Le séquençage des TALEs de 17.5 répétitions et plus, nécessite des runs supplémentaires de façon à amplifier les répétitions centrales à partir de nouvelles amorces (figure 11) et pouvoir ainsi compléter les données de RVDs. Le design de ces amorces est délicat car il faut cibler des SNPs (Single Nucleotide Polymorphism) présents sur une seule des répétitions. Plusieurs amorces ont été testées pour séquencer ces répétitions manquantes, toutes les amorces ne fonctionnant pas pour tous les TALEs. Il a donc fallu pour de nombreuses séquences dessiner des amorces au cas par cas, au plus près des régions manquantes (répétitions centrales), en s'assurant qu'elles ciblent un site unique dans la région répétée, ce qui n'est pas garanti: les séquences des régions répétées centrales étant inconnues, elles peuvent également contenir ces SNPs.



**Figure 12 : alignement des boxes théoriques cibles *pthA4-ptmA\*4* et *pthA\*2***

Le premier alignement est celui des séquences promotrices cibles de *pthA4* et *pthA\*4* montrant une forte homologie sur les 19 premières bases azotées. Les bases surlignées mettent en évidence les SNPs dans chaque séquence.

Le second alignement correspond aux séquences cibles du *pthA\*2*, c'est-à-dire non homologues à *pthA4*. Cet alignement montre que trois homologues chez des souches Aw présente de nombreux SNPs et donc peu d'homologie avec *pthA\*2*. A l'inverse, il met en évidence une homologie des *pthA\*2* les plus long avec des homologues de *pthA* de souches A.

### 3.1.2.3 Alignement des boxes promotrices

Les séquences promotrices cibles ont tout d'abord été déduites à l'aide du code établi par Bogdanove *et al.* 2010 et Boch *et al.* 2009 ; Streubel *et al.* 2012, les résultats sont présentés annexe 2. Toutes ces séquences ont été alignées à l'aide du logiciel Geneious à la recherche de motifs spécifiques d'un pathotype donné (figure 12 et 13). Les travaux de Streubel *et al.* ont par ailleurs montré l'importance des dirésidus "forts" : "HD" codant pour C, et "NN" codant pour G, et l'importance de leur répartition le long du TALE pour un bon fonctionnement (du moins lors d'essai sur TALEs artificiels).

#### ***a) Comparaison de l'homologue de 17.5 répétitions chez les souches du pathotype A à l'homologue de 18.5 répétitions des souches A\*.***

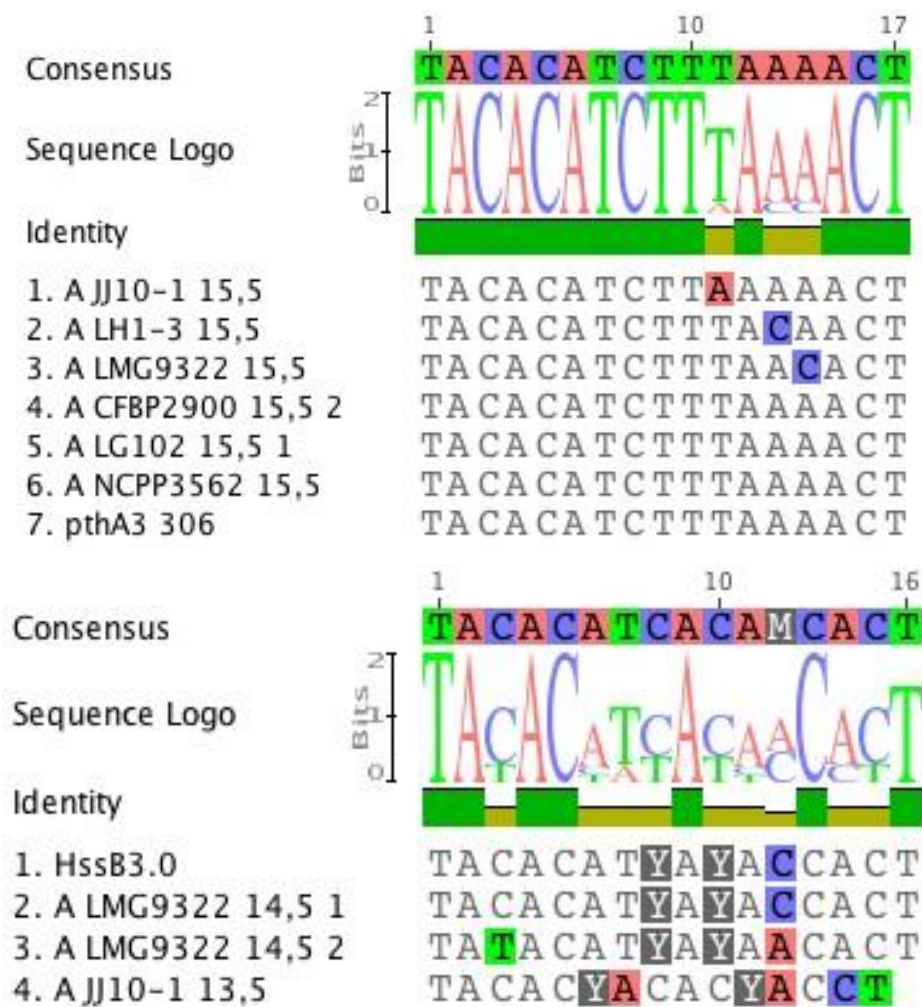
L'alignement de toutes les séquences théoriques cibles de *pthA\*4* (18.5 répétitions) à *pthA4* (17.5 répétitions), (figure 12) montre une forte homologie sur 19 bases, avec cependant une séquence sur deux présentant 1 à 2 SNPs, répartis sur 5 positions. L'analyse des RVDs correspondant aux SNPs montre pour 4 positions la variation entre RVD fort "HD" et faible "NI" ou "NG". On peut donc s'intéresser quand à l'importance de ces RVD dans la reconnaissance du promoteur.

Il est par conséquent probable que ces 2 TALEs activent les mêmes gènes et jouent le même rôle dans l'agressivité. Cela restant à confirmer par délétion de *pthA\*4* et recherche de phénotypes, pour voir si l'on obtient bien une perte de la capacité à faire du chancre comme après délétion de *pthA4* (Al-Saadi *et al.* 2007).

#### ***b) Analyse du TALE PthA\*2 = homologue de pthA chez les souches A\* dont la taille est différente de 18.5***

Chez les souches A\*, l'alignement de l'ensemble des TALEs non homologues à *pthA4*, c'est à dire ayant un nombre de répétitions différent de 18.5 (sauf pour la souche JF90-3) montre une séquence fortement conservée sur les 13 premières bases (figure 12), la fin variant avec les différences de longueurs. Par ailleurs, si on compare les homologues de 20.5 répétitions des souches A\* JK143-9 et JF90-2 aux TALEs de 20.5 répétitions des souches A C40 et JJ10-1, on remarque une très forte homologie, les homologues C40<sub>20.5</sub> et JJ10-1<sub>20.5</sub> étant possiblement 100% homologues (à une base prêt inconnue pour le moment du fait de runs de séquençage trop courts). cf figure 12. Ce TALE n'est donc a priori pas spécifique d'un pathotype donné. La délétion du *pthA\*2* a par ailleurs été réalisée chez la souche JF90-2 au laboratoire parallèlement à mon stage et le mutant correspondant JF90-2 $\Delta$ *pthA\*2* n'a pas





**Figure 13 : alignements de boxes promotrices cibles théoriques avec *pthA3* et *hssB3.0***

Le premier alignement montre les homologues des souches A présentant une forte homologie avec *pthA3*.

Le second alignement met en évidence les homologues des souches A ayant vraisemblablement les mêmes cibles que *hssB3.0*.



montré de différence du point de vue phénotypique en terme de date d'apparition des symptômes, d'agressivité et de gamme d'hôte (c'est-à-dire une gamme d'hôte restreinte au limettier et à *C. macrophylla*). Ce TALE n'a donc a priori pas de rôle dans la limitation de la gamme d'hôte par activation des gènes de résistances.

De plus, cet alignement a pu mettre en évidence trois *pthA*<sup>w</sup> dont les cibles semblent être différentes. Ces TALEs ont-ils le même rôle que *pthA*\*2 ? Activent t-ils les mêmes gènes ?

### **c) Autres homologues**

Chez les souches A, les séquences cibles des TALEs autres que celui de 17.5 répétitions ont été alignées. Le résultat a montré que les boxes cibles, pouvaient être classées en 5 groupes, avec chez la majorité des souches la présence d'un TALE ayant vraisemblablement la même cible que *pthA*3 (figure 13). Ce TALEs fortement conservé aurait-il un rôle dans l'élargissement de la gamme d'hôte ?

Enfin un alignement (figure 13) avec la séquence de *hssB3.0* (Shiotani *et al* 2007), a montré la présence d'homologue de ce gène responsable d'une restriction de gamme d'hôte, chez deux souches A (JJ10-1 et LMG9322). Ces souches présentant une forte diversité de TALE, on peut s'interroger sur la présence d'un autre homologue permettant de restaurer la gamme d'hôte.

### **3.1.2.4 Recherche des promoteurs cibles, *in silico***

Un extrait des résultats de prédictions des cibles des TALEs par Alvaro L. Pérez-Quintero dans le cadre d'une coopération avec l'IRD Montpellier, est présenté en annexe 3. Les résultats ont été triés en fonction d'un score qui prend en compte la ressemblance des boxes cibles dans le promoteur du gène de l'hôte aux boxes déduites du code à partir de la séquence des TALEs.

Un certain nombre de gènes impliqués dans les interactions plantes - pathogènes ressortent de ces analyses, notamment des gènes impliqués dans les phénomènes de résistances (protéines de type LRR (Leucine Rich Repeat), ou par exemple, des gènes habituellement activés par l'auxine (SAUR-like auxine responsive protein family), ce qui a pour effet de favoriser l'infection. Cependant, la plupart des cibles dévoilées par cette analyse sont des protéines peu connues dont l'implication dans les interactions plante-pathogène n'est pas ou peu décrite. La collaboration avec l'équipe de Montpellier permettra de déterminer plus précisément les cibles intéressantes à valider en RT-PCR. Il sera intéressant de vérifier si les TALEs de 17.5 et 18.5



ciblent les mêmes gènes, mais aussi de déterminer s'il est possible de mettre en évidence des cibles impliquées spécifiquement dans l'interaction entre une souche du pathotype A ou du pathotype A\* avec son hôte. Par exemple, il sera intéressant d'approfondir si le TALE *pthA3* fortement conservé chez les souches du pathotype A active des gènes spécifiques qui ne seraient pas activés par les TALEs présents chez les souches du pathotype A\*, ce qui pourrait expliquer la plus large gamme d'hôtes qui caractérise le pathotype A.

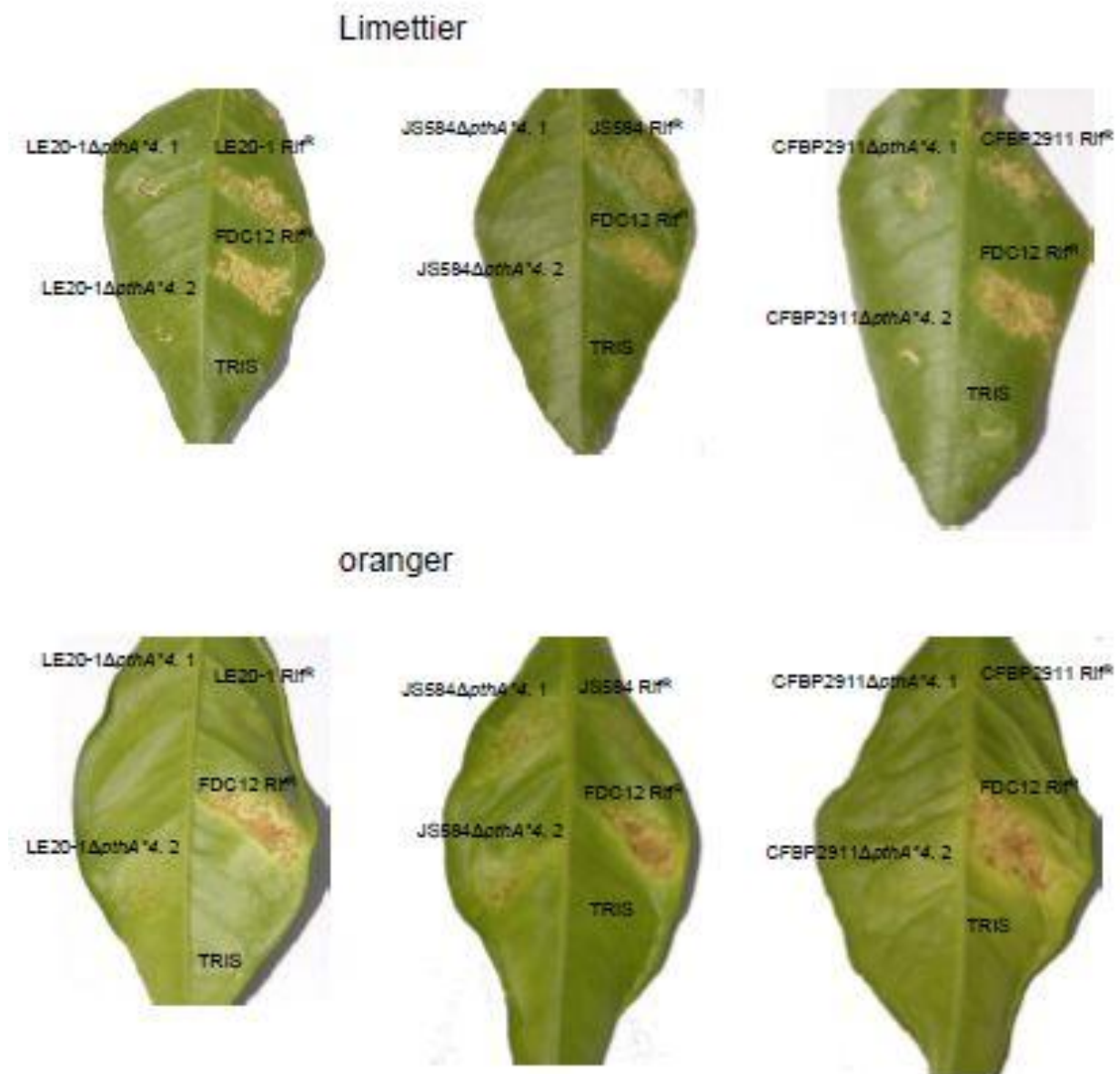
### **3.2 Déléction du gène *pthA\*4* chez 3 souches de *Xcc A\****

Les études préliminaires par Southern-blot et PCR sur la diversité des TALEs ont mis en évidence chez les souches A\* la présence d'un TALE conservé de 18.5 répétitions, ce qui ne correspond pas aux données de la littérature qui indiquent que l'ensemble des souches responsables du chancre des agrumes contiennent toutes un TALE de 17.5 répétitions responsable des symptômes d'hyperplasie et d'hypertrophie typiques du chancre (Al-Saadi *et al* 2007). Les résultats de clonage et séquençage ont confirmé qu'il s'agit bien de deux TALEs différents en terme de nombre de répétitions, mais fixant des boxes promotrices cibles théoriques très proches. Afin de vérifier que ces homologues de *pthA4* jouent bien un même rôle dans l'agressivité, le gène *pthA\*4* a été délété chez plusieurs souches A\*, puis les mutants ont été inoculés sur plantes.

#### **3.2.1 Amplification et clonage des régions flanquantes**

L'amplification des régions flanquantes du gène *pthA\*4* a été réalisée à l'aide d'amorces dessinées dans les régions flanquantes du gène *pthA4* de la souche de référence *Xcc* 306. La conservation de ces régions entre les différentes souches destinées à la délétion de *pthA\*4* étant inconnue, il a été décidé de construire un vecteur de délétion pour chacune des souches à muter, ceci afin de favoriser au maximum les phénomènes de recombinaison homologues lors de l'intégration du plasmide puis de la délétion du TALE, ceci grâce à une parfaite homologie de séquence.

Les régions flanquantes amplifiées et purifiées ont ensuite été clonées dans le vecteur pΔ13 selon le principe du golden-gate. Les plasmides obtenus, ont été introduits dans *E.coli* TG1 par électroporation, cette dernière jouant le rôle de bactérie donneuse dans la conjugaison tri-



**Figure 14 : Observation phénotypique sur limettier et oranger à J+21**

Photographies des symptômes sur les feuilles de la répétition A à J+21 après inoculation.

La première ligne montre les résultats obtenus sur Limettier. Tous les mutants montrent l'absence de pustules chancreuse, ainsi que le témoin TRIS. A l'inverse il y a une forte formation chancreuse pour la souche FDC12 (=IAPAR 306) pathogène sur tous les agrumes et pour les souches de références (LE20-1, JS584 et CFBP2911 Rif<sup>R</sup>) pathogènes sur Limettier.

La seconde ligne montre les symptômes obtenus sur oranger. La souche FDC12 confirme son pouvoir pathogène sur tous les agrumes par la formation de pustules chancreuse, alors que les souches de références A\*, elles, confirment leur gamme d'hôte restreinte en n'induisant pas de formation de chancre. Les mutants testés n'induisent pas non plus le chancre, cependant les souches JS584 (référence et mutants) ont induit la formation de pustules aqueuses, confirmant le caractère phénotypique atypique de cette souche observée préalablement.

parentale. Après sélection par PCR des clones contenant les régions flanquantes, deux clones sont conservés pour la suite des expérimentations, minipréparations pour un séquençage de contrôle et conjugaison tri-parentale

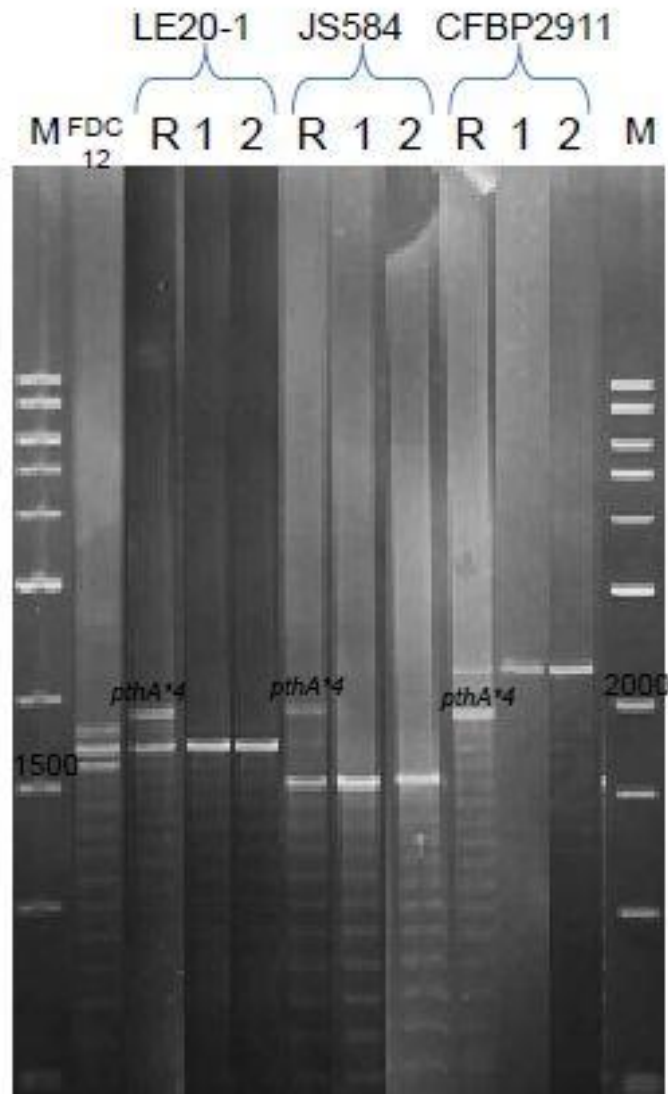
### 3.2.2 Résultat de la conjugaison tri-parentale

La conjugaison tri-parentale est une manipulation délicate qui donne des résultats variables, certaines souches étant plus facilement transformables que d'autres, sans qu'une explication claire ne soit donnée. Par conséquent les cinq souches à muter ont donné des résultats variables lors de la conjugaison triparentale.

Pour deux souches, JF90-2 et JF90-8, la première recombinaison (insertion du plasmide de délétion) n'a pas fonctionné, alors que la séquence du plasmide de délétion est correcte. La souche LE20-1 a donné le plus grand nombre de mutants, que se soit pour l'insertion du plasmide de délétion (première recombinaison) ou l'expulsion du gène (2e recombinaison). La sélection des mutants a été délicate pour la souche JS584, du fait de l'amplification non spécifique avec les amorces de vérification de la présence du plasmide (amorces E-MG); en effet, certaines bandes parasites n'ont pas pu être éliminées par modification du protocole de PCR et ont rendu la lecture du gel difficile. Deux clones ont tout de même été choisis pour la suite des expérimentations. Enfin la souche CFBP2911 n'a donné des mutants de délétion  $\Delta PthA^*4$  qu'à partir d'un seul clone obtenu lors de la première recombinaison contenant le plasmide de délétion.

### 3.2.3 Résultat de phénotypes sur plantes.

Les mutants obtenus ont été inoculés par infiltration sur des feuilles de limettier mexicain et sur oranger. Pour chaque souche A\* mutée, 2 mutants de délétion ont été comparés à la souche A\* correspondante (mutant Rif<sup>R</sup> pour faciliter les manipulations), ainsi qu'à un témoin positif (souche *Xcc* 306 induisant le chancre sur les 2 hôtes) et un témoin négatif (TRIS). Pour chaque souche 5 points d'inoculations ont été faits sur 5 feuilles différentes réparties sur 3 plants pour minimiser l'effet plante/feuille et appuyer les phénotypes observés. Les résultats sont présentés figure 14 et annexe 4: la souche 306 de pathotype A à large gamme d'hôtes a, comme attendu, induit la formation de chancre sur limettier et oranger dès 10 jours après inoculation. Les trois souches A\* ont induit la formation de chancre uniquement sur limettier mexicain 10 jours après inoculation, ce qui est en accord avec le phénotype des souches de pathotypes A\* à gamme d'hôtes restreinte.



**Figure 15 : vérification des génotypes bactériens sur symptômes**

Gel d'agarose après amplification des TALEs sur l'ADN extrait des bactéries, elles mêmes extraites des symptômes.

Le témoin FDC12 (souche *Xcc* 306) montre les trois tailles de TALEs attendues, dont celle correspondant à *pthA4* de 17.5 répétitions, montrant une fois de plus la différence de taille avec son homologue *pthA\*4*.

Pour les trois souches testées sur plantes (LE20-1 ; JS584 et CFBP2911) les résultats d'amplification de l'ADN respectivement extraits des souches de référence et des deux mutants testés, montrent bien l'absence du *pthA\*4* avec une conservation du second homologue, chez les mutants.

Sur limettier, les 3 mutants LE20-1 $\Delta pthA^*4$ , CFPB2911 $\Delta pthA^*4$  et JS584 $\Delta pthA^*4$  se sont avérés incapables d'induire les symptômes de type chancreux sur limettier 25 jours après inoculation, ce qui confirme le rôle de *pthA^\*4* dans l'induction des symptômes du chancre. En revanche, ces souches ont induit de petites taches aqueuses parfois accompagnées de points nécrotiques (figure 14), probablement témoins de réactions de défense de la plante.

Sur oranger, les mutants A\* LE20-1 $\Delta pthA^*4$ , CFPB2911 $\Delta pthA^*4$  se sont comportés comme la souche non mutée (respectivement LE20-1<sup>R</sup> et CFBP2911<sup>R</sup>), produisant de rares taches aqueuses. La délétion de *pthA^\*4* n'a donc pas modifié la réaction non hôte pour ces souches. Le mutant JS584 $\Delta pthA^*4$  a également produit des taches aqueuses, alors que la souche JS584<sup>R</sup> a induit de petites pustules, ce qui est conforme aux résultats précédents obtenus dans l'équipe, qui montrent que cette souche est caractérisée par une agressivité intermédiaire entre les A et les A\*.

Cette expérimentation confirme que l'homologue de *pthA4* qui contient 18.5 répétitions chez les 3 souches A\* étudiées joue le même rôle que l'homologue de 17.5 répétitions chez les souches du pathotype A: il est indispensable à la formation des symptômes d'hyperplasie et hypertrophie cellulaire caractéristiques du chancre (Al-Saadi *et al.* 2007). Ce résultat reste à confirmer par des tests de complémentation, une restauration de l'agressivité permettant d'affirmer ce rôle.

### 3.2.4 Vérification des géotypes par PCR sur symptômes

Afin de vérifier le géotype et la mutation des bactéries responsables des phénotypes observés, celles-ci ainsi que les contrôles (souches sauvages Rif<sup>R</sup>), ont été extraites à partir des symptômes sur limettier. L'absence de *pthA^\** chez les mutants et sa présence chez les souches contrôle avant et après inoculation a été confirmée par PCR en utilisant les couples d'amorces E-3R et E-MG (figure 15). Les profils obtenus par PCR confirment l'absence du TALE de 18.5 répétitions pour les 3 souches mutées. Un contrôle par Southern-blot, technique plus sensible, permettrait de confirmer de manière non équivoque le contenu en TALE de ces souches mutantes et confirmer les observations des phénotypes sur plante.





## 4 Conclusion et perspectives:

Le séquençage de l'ensemble des TALEs d'une collection de souches représentatives de la diversité de *Xcc* en terme de pathotypes et d'origines géographiques a permis de mettre en évidence plusieurs points intéressants, ouvrant de nouvelles perspectives de travail pour comprendre l'implication des TALEs dans la spécialisation d'hôte chez *Xcc*.

Le séquençage systématique de plusieurs homologues pour chaque taille de TALE a confirmé une plus grande diversité chez les souches de pathotype A par rapport aux souches du pathotype A\*, les souches du pathotype A possédant un plus grand nombre de TALEs avec une plus grande diversité en terme de taille et de séquence (séquence des RVDs). Les études de diversité de *Xcc* précédemment réalisées dans l'équipe à l'aide de différents marqueurs moléculaires ont mis en évidence une plus grande diversité génétique pour les souches du pathotype A\* par rapport aux souches du pathotype A. Et donc, la diversité qui caractérise les TALEs des souches du pathotype A ne coïncide pas avec les données de diversité génétique. Est-ce cette diversité qui permet aux souches A de contourner les mécanismes de défenses de la plante ? Les travaux de Shiotani *et al* en 2007 ont montré des phénomènes de recombinaison entre les TALEs aboutissant à des modifications de la gamme d'hôte pour un petit groupe de souches présentant un phénotype particulier sur pamplemoussier Tachibana. Il est probable que la diversité des TALEs chez les souches du pathotype A constitue un réservoir permettant des événements de recombinaison et une adaptation à un nouvel hôte.

Notre étude a mis en évidence lors de l'analyse du *pthA*\*2, une conservation des 13 premiers RVDs. Des homologues proches de ces TALEs ont été retrouvés chez quelques souches du pathotype A, ce qui indique qu'*a priori* ce TALE n'est pas impliqué dans la définition de la gamme d'hôtes. Des tests de délétion de ce TALE ont été effectués parallèlement à mon projet et la délétion de *pthA*\*2 chez les souches A\* n'a pas entraîné de modification de la gamme d'hôte.

L'analyse des séquences de RVDs de *pthA*4 et *pthA*\*4, et des boîtes promotrices cibles déduites du code établi par Bogdanove *et al* 2009 et Boch *et al* 2009, a montré une forte homologie entre ces TALEs de tailles différentes, respectivement 17.5 et 18.5 répétitions, et la délétion de *pthA*\*4 a confirmé que chez les souches du pathotype A\*, le TALE responsable



des symptômes du chancre a 18.5 répétitions et non 17.5 répétitions comme indiqué dans la littérature.

Enfin, un résultat particulièrement intéressant ressort de ce travail: il existe un homologue de *pthA3* chez l'ensemble des souches du pathotype A incluses à cette étude. Des travaux de délétion sont prévus afin de déterminer le rôle de ce TALE dans la détermination de la gamme d'hôtes chez *Xcc*.

Le travail de mise au point effectué a permis de mettre en place un certain nombre de protocoles pour la manipulation des TALEs de *Xcc*. Au regard de mes résultats, des travaux de délétion sont prévus afin de valider certaines hypothèses et afin de déterminer si chez *Xcc*, du fait du potentiel évolutif fort qui les caractérise, les TALEs participent à l'adaptation du pathogène à son hôte. Parallèlement, l'investigation de la diversité des TALEs chez *Xcc* vient en complément des travaux de compréhension du fonctionnement des TALEs réalisés dans les équipes se focalisant principalement sur TALEs artificiels (Boch *et al* 2009, Streubel *et al* 2012), notamment pour répondre à des questions de tolérance de certains SNPs dans l'activation des gènes cibles de la plante hôte.



## **Bibliographie**

- Ah-You, N., Gagnevin, L., Grimont, P. A. D., Brisse, S., Nesme, X., Chiroleu, F., *et al.* (2008) Polyphasic characterization of xanthomonads pathogenic to Anacardiaceae and their relatedness to different *Xanthomonas* species. *Int. J. Syst. Evol. Microbiol.*, **58**, 000-000.
- Al-Saadi, A. et al., 2007. All five host-range variants of *Xanthomonas citri* carry one pthA homolog with 17.5 repeats that determines pathogenicity on citrus, but none determine host-range variation. *Molecular plant-microbe interactions : MPMI*, 20(8), pp.934–43.
- Boch, J. et al., 2009. Breaking the code of DNA binding specificity of TAL-type III effectors. *Science (New York, N.Y.)*, 326(5959), pp.1509–12. Bogdanove, A.J., Schornack, S. & Lahaye, T., 2010. TAL effectors: finding plant genes for disease and defense. *Current opinion in plant biology*, 13(4), pp.394–401.
- Brunings, Asha M ; Gabriel, D.W., 2003. *Xanthomonas citri*: breaking the surface. *Molecular plant pathology*, 4, pp.141–157.
- Bui, T.N.L. et al., 2009. From local surveys to global surveillance: three high-throughput genotyping methods for epidemiological monitoring of *Xanthomonas citri* pv. *citri* pathotypes. *Applied and environmental microbiology*, 75(4), pp.1173–84.
- Civerolo, E., 1984. Bacterial canker disease of citrus. *Journal of the Rio Grande Valley Horticultural Society*, 37, pp.127–146.
- Das, A., 2003. Citrus canker-A review. *J. Appl. Hort*, 5(1), pp.52–60. Available at: [http://horticultureresearch.net/Citrus canker.pdf](http://horticultureresearch.net/Citrus%20canker.pdf) [Accessed August 9, 2012].
- Deng, D. et al., 2012. Structural basis for sequence-specific recognition of DNA by TAL effectors. *Science (New York, N.Y.)*, 335(6069), pp.720–3.
- Doidge, E.M., 1916. Citrus canker in South Africa. *South African Fruit Grower*, 3, pp.265–268.
- Gottwald, Tim R, Pierce, F. & Graham, J.H., 2002. Citrus Canker : The Pathogen and Its Impact Plant Health Progress Plant Health Progress. *Plant Health Progress*.
- Graham, J.H. et al., 2004. *Xanthomonas axonopodis* pv . *citri* : factors affecting successful. *Molecular plant pathology*, 5, pp.1–15.
- Gürlebeck, D., Thieme, F. & Bonas, U., 2006. Type III effector proteins from the plant pathogen *Xanthomonas* and their role in the interaction with the host plant. *Journal of plant physiology*, 163(3), pp.233–55.
- Lee, S., 2008. No TitleDiversity of pthA gene of *Xanthomonas* Strains Causing citrus Bacterial Canker and its relationship with virulence. *Plant Pathology*, 24, pp.357–360.



- Mak, A.N.-S. et al., 2012. The crystal structure of TAL effector PthXo1 bound to its DNA target. *Science (New York, N.Y.)*, 335(6069), pp.716–9. Rossetti, V., 1977. Citrus Canker in latin America: a review. *Proc.Int.Soc.citriculture*.
- Shiotani, H. et al., 2007. A pthA homolog from *Xanthomonas axonopodis* pv. citri responsible for host-specific suppression of virulence. *Journal of bacteriology*, 189(8), pp.3271–9. da Silva, a C.R. et al., 2002. Comparison of the genomes of two *Xanthomonas* pathogens with differing host specificities. *Nature*, 417(6887), pp.459–63.
- Streubel, J. et al., 2012. TAL effector RVD specificities and efficiencies. *Nature biotechnology*, 30(7), pp.593–5.
- Sun, X., Stall, R. & Jones, J., 2004. Detection and characterization of a new strain of citrus canker bacteria from Key/Mexican lime and alemow in South Florida. *Plant disease*, 88(November), pp.1179–1188.
- Swarup, S. et al., 1991. A Pathogenicity Locus from *Xanthomonas citi*T; Enables Strains from Severa) Pathovars of. *Molecular plant pathology*, 81, pp.802–809.
- Swarup, S. et al., 1992. An *Xanthomonas citri* pathogenicity gene, pthA, pleiotropically encodes gratuitous Avirulence on nonhosts. *Molecular plant-microbe interactions : MPMI*, 5(3), pp.204–213.





# Annexes



## Annexe 1 : Protocole de Minipréparations

### Wizard® Plus SV Minipreps DNA Purification System

INSTRUCTIONS FOR USE OF PRODUCTS A1330, A1340, A1460, A1465 AND A1470.

**Quick**  
PROTOCOL

#### Centrifugation Protocol

##### Production of Cleared Lysate

1. Pellet 1–10ml of overnight culture for 5 minutes.
2. Thoroughly resuspend pellet with 250µl of Cell Resuspension Solution.
3. Add 250µl of Cell Lysis Solution to each sample; invert 4 times to mix.
4. Add 10µl of Alkaline Protease Solution; invert 4 times to mix. Incubate 5 minutes at room temperature.
5. Add 350µl of Neutralization Solution; invert 4 times to mix.
6. Centrifuge at top speed for 10 minutes at room temperature.

##### Binding of Plasmid DNA

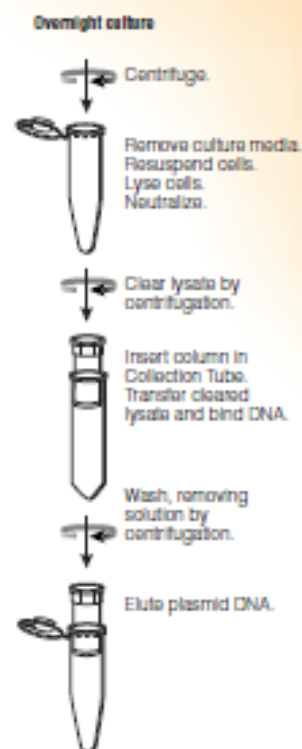
7. Insert Spin Column into Collection Tube.
8. Decant cleared lysate into Spin Column.
9. Centrifuge at top speed for 1 minute at room temperature. Discard flowthrough, and reinsert Column into Collection Tube.

##### Washing

10. Add 750µl of Wash Solution (ethanol added). Centrifuge at top speed for 1 minute. Discard flowthrough and reinsert column into Collection Tube.
11. Repeat Step 10 with 250µl of Wash Solution.
12. Centrifuge at top speed for 2 minutes at room temperature.

##### Elution

13. Transfer Spin Column to a sterile 1.5ml microcentrifuge tube, being careful not to transfer any of the Column Wash Solution with the Spin Column. If the Spin Column has Column Wash Solution associated with it, centrifuge again for 1 minute at top speed, then transfer the Spin Column to a new, sterile 1.5ml microcentrifuge tube.
14. Add 100µl of Nuclease-Free Water to the Spin Column. Centrifuge at top speed for 1 minute at room temperature.
15. Discard column, and store DNA at –20°C or below.



1591 M039\_0A

Additional protocol information is available in Technical Bulletin #TB225, available online at: [www.promega.com](http://www.promega.com)

#### ORDERING/TECHNICAL INFORMATION:

[www.promega.com](http://www.promega.com) • Phone 608-274-4330 • Fax 608-277-2601

©1999, 2001, 2002, 2005, 2009 Promega Corporation. All Rights Reserved.



**Promega**

Printed in USA. Revised 12/09  
Part #3FB004



## Annexe 2 : Relevé des RVDs et détermination des boxes promotrices cibles

pv	Souche	nb essais	nb rep	séquence des RVDs	box promotrice cible	remarques
A	LG102	1	15.5	NI-HD-NI-HD-NI-NG-HD-NG-NG-NG-NI-NI-NI-NI-HD-NG	TACACATCTTTAAAACT	
		3	15.5	NI-N.-NI-N.-NI-HD-HD-N.-NI-HD-NI-HD-N.-NI-HD-NG	TAYAYACCYACACYACT	
		1	17.5	NI-N.-NG-N.-NI-HD-HD-NG-HD-NG-NG-NG-NG-NS-HD-HD-NG-NG	TAYTYACCTCTTTTNCCTT	artefact
		3	17.5	NI-N.-NI-HD-NI-HD-HD-NG-HD-NG-NG-NG-NG-NS-HD-HD-NG-NG	TAYACACCTCTTTTNCCTT	
		4	19.5	NI-NI-NG-HD-HD-NI-NI-NG-NG-NG-HD-NI-NG-HD-NG-NG-NG-NG-NI-NG	TAATCCAATTTTCATCTTTTAT	
		1	23.5	NI-HD-NS-NI-HD-NS-NI-NS-NS-NI-HD-NI-NG-NI-NI-NS-HD-HD-NS-HD-HD-NS-NG-NG	TACNACNANNACATAANCCNCCNTT	
		1	22.5	NI-HD-NS-NI-HD-NS-NI-NS-NS-NS-NI-HD-NI-NS-HD-HD-NS-HD-HD-NS-NG-NG	TACNACNANANNACANCCNCCNTT	artefact
	LMG9322	1	14.5	NI-HD-NI-HD-NI-NG-N.-NI-N.-NI-HD-HD-NI-HD-NG	TACACATYAYACCACT	
		6	14.5	NI-NG-NI-HD-NI-NG-N.-NI-N.-NI-NI-HD-NI-HD-NG	TATACATYAYAACACT	
		4	15.5	NI-HD-NI-HD-NI-NG-HD-NG-NG-NG-NG-NI-NI-HD-NI-HD-NG	TACACATCTTTAAACACT	
		5	16.5	NI-N.-NI-N.-NI-HD-HD-N.-NI-HD-NI-HD-N.-NI-HD-HD-NG	TAYAYACCYACACYACCT	
		5	17.5	NI-N.-NI-NI-NI-NG-HD-NG-HD-NG-NG-NG-NG-NS-HD-HD-NG-NG	TAYAAATCTCTTTTNCCTT	
	JJ10-1	2	8.5	NI-HD-NI-NI-NI-NI-NI-HD-NG	TACAAAAACT	
		3	9.5	NI-HD-NI-NI-HD-NI-HD-NI-HD-NG	TACAACACACT	
		4	13.5	NI-HD-NI-HD-N.-NI-HD-NI-HD-N.-NI-HD-HD-NG	TACACYACACYACCT	
		5	15.5	NI-HD-NI-HD-NI-NG-HD-NG-NG-NG-NI-NI-NI-NI-NI-HD-NG	TACACATCTTAAAAACT	
		4	17.5	NI-N.-NI-NI-NI-HD-HD-NG-HD-NG-NG-NG-NG-NS-HD-HD-NG-NG	TAYAAACCTCTTTTNCCTT	
		3	20.5	NI-N.-NI-N.-NI-HD-HD-N.-HD-NI-HD-N.-NI-HD-NI-HD-N.-NI-HD-HD-NG	TAYAYACCYACACYACCT	
	CFBP2900	4	14.5	NI-HD-NI-HD-N.-NI-HD-HD-NI-HD-N.-NI-HD-HD-NG	TACACNACCACNACCT	
		1	15.5	NI-N.-NI-NI-HD-N.-NI-HD-HD-NI-HD-N.-NI-HD-HD-NG	TAYAACYACCACYACCT	
		3	15.5	NI-HD-NI-HD-NI-NG-HD-NG-NG-NG-NG-NI-NI-NI-NI-HD-NG	TACACATCTTTAAAACT	
		5	17.5	NI-N.-NI-NI-NI-HD-HD-NG-HD-NG-NG-NG-NG-NS-HD-HD-NG-NG	TAYAAACCTCTTTTNCCTT	
		4	21.5	NI-N.-NI-N.-NI-HD-HD-N.-NI-HD-NI-12-NI-HD-HD-NI-HD-N.-NI-HD-HD-NG	TAYAYACCYACAXACCACYACCT	incomplete
	LH1-3	4	15.5	NI-HD-NI-HD-NI-NG-HD-NG-NG-NG-NG-NI-HD-NI-NI-HD-NG	TACACATCTTTACAACCT	
		1	17.5	NI-NG-NG-NG-NS-HD-HD-NG-HD-NG-NG-NG-NG-NS-HD-HD-NG-NG	TATTTNCCTCTTTTNCCTT	
		1	18.5	NI-N.-NI-N.-NI-N.-NI-HD-HD-N.-NI-HD-NI-HD-N.-HD-NG-HD-NG	TAYAYAYACCYACACYCTCT	artefact
		1	19.5	NI-N.-NI-N.-NI-N.-NI-N.-NI-HD-11-12-HD-NI-HD-N.-HD-NG-HD-NG	TAYAYAYAYACXXCACYCTCT	incomplete
	NCP3562	3	15.5	NI-HD-NI-HD-NI-NG-HD-NG-NG-NG-NG-NI-NI-NI-NI-HD-NG	TACACATCTTTAAAACT	
		2	17.5	NI-NG-NG-NG-NS-HD-HD-NG-HD-NG-NG-NG-NG-NS-HD-HD-NG-NG	TATTTNCCTCTTTTNCCTT	
		3	18.5	NI-N.-NI-N.-NI-N.-NI-HD-HD-N.-NI-HD-NI-HD-N.-HD-NG-HD-NG	TAYAYAYACCYACACYCTCT	
	NCP3610	2	11.5	NI-NG-NI-HD-NI-HD-N.-NI-NI-HD-HD-NG	TATACACYAACCT	
		3	12.5	NI-NI-NG-HD-NG-NG-NG-NG-NI-NG-NI-NI-HD-NG	TAATCTTTATAACT	
			17.5			
		2	20.5	NI-N.-NI-N.-NI-HD-HD-N.-NI-HD-11-12-13-HD-NI-HD-N.-NI-HD-HD-NG	TAYAYACCYACXXXCACYACCT	incomplete
	C40	3	20.5	NI-N.-NI-N.-NI-HD-HD-N.-NI-HD-NI-N.-NI-HD-NI-HD-N.-NI-HD-HD-NG	TAYAYACCYACAYACACYACCT	



## Annexe 2 : Relevé des RVDs et détermination des boxes promotrices cibles

pv	souche	nb essais	nb rep	séquence des RVDs	box promotrice cible	remarques
A*	JJ238-24	3	18.5	NI-N.-NI-HD-NI-HD-HD-NG-HD-NG-NG-NG-NG-NS-HD-NS-NG-NG-NG	TAYACACCTCTTTTNCNTTT	
			20.5			
	JS555	4	14.5	NI-N.-NI-N.-NI-HD-HD-NG-NI-HD-NI-HD-HD-HD-NG	TAYAYACCTACACCCT	
			18.5			
	JS581	2	13.5	NI-N.-NI-N.-NI-HD-HD-NG-NI-HD-NI-HD-HD-NG	TAYAYACCTACACCCT	
		1	18.5	NI-N.-NI-HD-NI-HD-NG-NG-HD-NG-NG-NG-NG-NS-HD-NS-NG-NG-NG	TAYACACTCTTTTNCNTTT	
	LE3-1	4	16.5	NI-N.-NI-N.-NI-HD-HD-N.-NI-HD-NI-HD-HD-NI-N.-HD-NG	TAYAYACCYACACCAYCT	
		2	18.5	NI-N.-NI-HD-NI-HD-HD-NG-HD-10-NG-NG-NG-NS-HD-NS-NG-NG-NG	TAYACACCTCXTTTNCNTTT	incomplete
	CFBP2911	4	18.5	NI-N.-NI-HD-NI-HD-HD-NG-HD-NG-NG-NG-NG-NS-HD-NS-NG-NG-NG	TAYACACCTCTTTTNCNTTT	
		3	21.5	NI-N.-NI-N.-NI-HD-HD-NG-NI-HD-NI-HD-N.-NI-HD-NI-HD-N.-NI-HD-HD-NG	TAYAYACCTACACYACACYACCT	
	JF90-2	3	18.5	NI-N.-NI-HD-NI-HD-HD-NG-HD-NG-NG-NG-NG-NS-HD-NS-NG-NG-NG	TAYACACCTCTTTTNCNTTT	
		3	21.5	NI-N.-NI-N.-NI-HD-HD-NG-NI-HD-NI-HD-N.-NI-HD-NI-HD-N.-NI-HD-HD-NG	TAYAYACCTACACYACACYACCT	
	JK143-9	5	18.5	NI-N.-NI-HD-NI-HD-HD-NG-HD-NG-NG-NG-NG-NS-HD-NS-NG-NG-NG	TAYACACCTCTTTTNCNTTT	
		4	21.5	NI-N.-NI-N.-NI-HD-HD-N.-NI-HD-HD-HD-HD-N.-NI-N.-N.-NI-HD-HD-NG	TAYAYACCYACCCCCYAYACCT	
	JK48	3	14.5	NI-N.-NI-N.-NI-HD-HD-NG-NI-HD-NI-HD-HD-HD-NG	TAYAYACCTACACCCT	
		3	18.5	NI-NG-NI-HD-NI-HD-HD-NG-HD-NG-NG-NG-NG-NS-HD-NS-NG-NG-NG	TATACACCTCTTTTNCNTTT	
	90-3		18.5	NI-N.-NI-N.-NI-HD-HD-N.-NI-HD-NI-HD-HD-NI-HD-NG-HD-HD-NG	TAYAYACCYACCACTCCT	
			18.5	NI-NG-NI-HD-NI-HD-HD-NG-HD-NG-NG-NG-NG-NS-HD-NS-NG-NG-NG	TAYACACCTCTTTTNCNTTT	
A <sup>w</sup>	NCPB3608		18.5	NI-NG-NG-NG-NS-HD-HD-HD-HD-NG-NG-NG-NG-NS-HD-HD-NG-NG-NG	TATTTNCCCCTTTTNCCTTT	
			19.5	NI-N.-NI-NI-HD-NI-HD-HD-HD-NI-N.-NI-NS-NI-HD-NI-N.-NI-NS-NG	TAYAACACCCAYANACAYANT	
	JF90-8	4	14.5	NI-N.-NI-N.-NI-HD-HD-N.-NI-HD-NI-HD-N.-HD-NG	TAYAYACCYACACYCT	
		4	18.5	NI-N.-NI-HD-NI-HD-HD-NG-HD-NG-NG-NG-NG-NS-HD-NS-NG-NG-NG	TAYACACCTCTTTTNCNTTT	
	LG100	3	14.5	NI-N.-NI-N.-NI-HD-HD-N.-NI-HD-NI-HD-N.-HD-NG	TAYAYACCYACACYCT	
		3	18.5	NI-N.-NI-HD-NI-HD-HD-NG-HD-NG-NG-NG-NG-NS-HD-NS-NG-NG-NG	TAYACACCTCTTTTNCNTTT	
	X2003-3218	3	17.5	NI-NG-NG-D/NG-NS-HD-HD-NS-09-NG-NG-NG-NG-NS-HD-HD-NG-NG	TATTTNCCNXXTTTNCCTT	
			19.5	NI-N.-NI-NI-HD-NI-HD-HD-HD-10-11-HD-NS-NI-HD-NI-N.-NI-NS-NG	TAYAACACCCXXCNACAYANT	incomplete
	LG115	1	18.5	NI-NG-NI-HD-NI-HD-HD-NG-HD-NG-NG-NG-NG-NS-HD-NS-NG-NG-NG	TATACACCTCTTTTNCNTTT	
		3	19.5	NI-NG-NG-NG-NS-HD-HD-NS-HD-NG-NG-12-NG-NG-NS-HD-HD-NG-NG-NG	TATTTNCCNCTTXXTNCCTTT	incomplete





### **Annexe 3:** Extrait des boxes promotrices cibles prédites *in silico* pour un *pthA* de souche A

SEQ_ID	SCORE	DISTANCE_ FROM_END	INI	END	SEQ	RANK	ANNOTATION
orange1.1g026556m	15.604	-112	888	907.5	TATAAACCCCTTTGCCTT	1	LOB domain-containing protein 1
orange1.1g027107m	14.387	-238	762	781.5	TACAAGCCCTTCTCCCTC	2	Integrase-type DNA-binding superfamily protein
orange1.1g030516m	13.939	-178	822	841.5	CCGCAACCTCCTTAcctG	3	Adenine nucleotide alpha hydrolases-like superfamily protein
orange1.1g035371m	13.933	-763	237	256.5	CACCCACCTCTTTGCATG	4	0
orange1.1g013261m	13.932	-590	410	429.5	TACAAATCTCTCTCCCTG	5	sequence-specific DNA binding transcription factors
orange1.1g014072m	13.836	-807	193	212.5	CATCAACCTCTTTtACCAC	6	B-box type zinc finger protein with CCT domain
orange1.1g000742m	13.789	-73	927	946.5	CAAAACCTCTTTCCCCG	7	tetratricopeptide repeat (TPR)-containing protein
orange1.1g012054m	13.677	-237	763	782.5	TCCAAACACCTTtCCtC	8	bromodomain 4
orange1.1g043004m	13.663	-147	853	872.5	CCcACACCTCctCtACCTA	9	GNS1/SUR4 membrane protein family
orange1.1g031728m	13.634	-766	136	155.5	TATAAACCCCTTTCCAC	10	Poly (ADP-ribose) glycohydrolase (PARG)
orange1.1g028855m	13.606	-426	574	593.5	TAGGCACCTCTTTACCTT	11	Uncharacterised protein family (UPF0497)
orange1.1g034153m	13.523	-108	892	911.5	CCCAAACCTCtCTTCACTC	12	Preprotein translocase Sec, Sec61-beta subunit protein
orange1.1g010561m	13.512	-246	754	773.5	CAAAAACCTCTCCGCCTA	13	MATE efflux family protein
orange1.1g012188m	13.409	-512	488	507.5	CACCAACCTTTTTGCCTT	14	alpha/beta-Hydrolases superfamily protein
orange1.1g039048m	13.279	-911	89	108.5	CACACACCCCTCTTCTC	15	Argonaute family protein
orange1.1g044126m	13.262	-644	356	375.5	CCAAAACCTCTCCACCTG	16	Subtilase family protein
orange1.1g043792m	13.216	-115	885	904.5	TACAaCCCCCCCcTCCCTC	17	NAD(P)-binding Rossmann-fold superfamily protein
orange1.1g007269m	13.208	-430	570	589.5	CCCAAACCTCTTTAACCT	18	Rhamnogalacturonate lyase family protein
orange1.1g042813m	13.045	-755	245	264.5	CCGACACCTCCTTGCCCG	19	LOB domain-containing protein 20
orange1.1g004377m	12.646	-339	661	680.5	CACAAACCACTCTCCCTC	20	Enhancer of polycomb-like transcription factor protein
orange1.1g037682m	12.547	-523	477	496.5	CACCAACCAATTTACCTC	21	0
orange1.1g035959m	12.52	-119	23	42.5	CCAAAACCTCCTTTACCTT	22	cytochrome BC1 synthesis
orange1.1g031831m	12.514	-832	168	187.5	CATAAGCTTCTTTGCCTG	23	0



### **Annexe 3:** Extrait des boxes promotrices cibles prédites *in silico* pour un *pthA* de souche A

orange1.1g038696m	12.502	-400	600	619.5	CCCCCCCCcTtTGCCTG	24	Protein kinase family protein
orange1.1g005422m	12.477	-59	941	960.5	TATAGAACCCTTTTGCCTG	25	cyclic nucleotide-gated channel 15
orange1.1g034567m	12.441	-667	333	352.5	TATCAACCCCTTCTCCTC	26	0
orange1.1g021359m	12.434	-200	800	819.5	CACAAACCCcTCCCCcA	27	0
orange1.1g009663m	12.418	-479	521	540.5	CACAGACCCCTCCCCCTC	28	Pentatricopeptide repeat (PPR) superfamily protein
orange1.1g006859m	12.402	-669	331	350.5	TGTAGACCTTTTACCTT	29	MUTS-like protein 4
orange1.1g018215m	12.19	-540	460	479.5	CacAAACCCCCCCCCCCC	30	WRKY DNA-binding protein 7
orange1.1g002895m	12.158	-317	683	702.5	CCCAAACcTCCCCCCCCC	31	SET domain-containing protein
orange1.1g016384m	12.131	-777	223	242.5	CACAAACCTCTCCTCTCT	32	protein kinase C-like zinc finger protein
orange1.1g001858m	12.095	-525	475	494.5	TGCCCACCACTTTTCCCTG	33	Leucine-rich repeat protein kinase family protein
orange1.1g019645m	12.066	-53	947	966.5	CACCACCCTcTTTTCTTC	34	chloroplast outer envelope protein 37
orange1.1g035720m	12.049	-398	602	621.5	CCTAAACCTCTTTTACTT	35	Ribosomal protein L10 family protein
orange1.1g032877m	12.024	-90	910	929.5	TATAAACCTCCTTCCCAT	36	0
orange1.1g028490m	11.959	-800	200	219.5	TATACACCTTTCTTCCCTC	37	0
orange1.1g041802m	11.952	-236	764	783.5	CACCAACCCCTTTACCCA	38	ovate family protein 13
orange1.1g000806m	11.887	-394	606	625.5	CACACACcCctCTctCTC	39	0
orange1.1g006968m	11.887	-836	164	183.5	CACACCCCTCTTACCCTT	40	NSP-interacting kinase 3
orange1.1g030886m	11.857	-570	430	449.5	TACAACCCcTCCCCcAC	41	0
orange1.1g009793m	11.843	-120	880	899.5	CATAACCCCATTTCCCTT	42	cytokinin oxidase 3
orange1.1g005944m	11.837	-644	356	375.5	CATaAAACTCTATTACCTA	43	NPK1-related protein kinase 3
orange1.1g030559m	11.802	-877	123	142.5	TCCAGCCCTCTTCCCT	44	TATA binding protein associated factor 21kDa subunit
orange1.1g015442m	11.769	-215	785	804.5	TCCGCCCTCTTTCTCTC	45	Transcription factor IIA, alpha/beta subunit
orange1.1g038699m	11.736	-832	168	187.5	TCCAAACCTCCTTCGCTT	46	Leucine-rich repeat receptor-like protein kinase family protein



### **Annexe 3:** Extrait des boxes promotrices cibles prédites *in silico* pour un *pthA* de souche A

orange1.1g022744m	11.72	-954	46	65.5	TCCAAACCTCTCTTACTCT	47	UDP-glucosyltransferase 74F2
orange1.1g046358m	11.709	-965	35	54.5	CACAAACATCATTTCCT	48	Pollen Ole e 1 allergen and extensin family protein
orange1.1g038571m	11.696	-37	963	982.5	CCCAAACTCTTCTGCCAG	49	Homeodomain-like superfamily protein
orange1.1g035762m	11.692	-285	715	734.5	TATAAGCCTTTTTaCCCC	50	0
orange1.1g025989m	11.642	-215	785	804.5	CCGAACCTCATTGCCCC	51	differentiation and greening-like 1
orange1.1g006781m	11.635	-117	883	902.5	TAAAAGCCTCCTTCCTC	52	with no lysine (K) kinase 8
orange1.1g046419m	11.632	-802	198	217.5	CCTAAaCCCaTCTCCctc	53	myb domain protein 97



## Annexe 4: relevé des symptômes sur limettier et oranger

### Souches :

1 : LE20-1      3 : CFBP2911

2 : JS584      20 : FDC12

### Légende :

rP= rares pustules

P= pustules

Pa= pustules aqueuses

Pc= pustules chancreuses

J+15

DATE: 18/06/12

Feuille	Souche	Oranger					Limettier				
		A	B	C	D	E	A	B	C	D	E
1	1	rP	-	-	-	-	-	-	-	-	-
1	2	rP	-	-	-	-	-	-	-	-	-
1	3	rP	-	-	-	-	Pc	Pc+	Pc	P	Pc
1	20	Pg	Pc	P	Pc	Pc	Pc+	Pc	Pc	P+	Pc
1	TRIS	-	-	-	-	-	-	-	-	-	-
2	4	P	P	rP	rP	P	-	-	-	-	-
2	5	P	P	rP	rP	P	-	-	-	-	-
2	6	P	P	rP	rP	P	P	*P	P	-	-
2	20	Pc	Pc	Pc	Pc+	Pc+	Pc	Pc	Pc	Pc	-
2	TRIS	-	-	-	-	-	-	-	-	-	-
3	7	-	-	-	-	-	-	-	-	-	-
3	8	-	-	-	-	-	-	-	-	-	-
3	9	-	-	-	-	-	rP	rP	rP	rP	rP
3	20	Pc	Pc	Pc	Pc	Pc+	Pc	Pc	Pc	Pc	Pc
3	TRIS	-	-	-	-	-	-	-	-	-	-

J+25

DATE: 28/06/12

Feuille	Souche	Oranger					Limettier				
		A	B	C	D	E	A	B	C	D	E
1	1	Pg	Pg+	Pg	Pg	-	Pg	Pg	Pg	Pg	-
1	2	Pg	Pg+	Pg	Pg	-	Pg	Pg	Pg	Pg	-
1	3	Pg	Pg+	Pg	Pg	-	Pg	Pg	Pg	Pg	-
1	20	Pg	Pg+	Pg	Pg	-	Pg	Pg	Pg	Pg	-
1	TRIS	-	-	-	-	-	-	-	-	-	-
2	4	Pg	Pg+	Pg	Pg	-	Pg	Pg	Pg	-	-
2	5	Pg	Pg+	Pg	Pg	-	Pg	Pg	Pg	-	-
2	6	Pg	Pg+	Pg	Pg	-	Pg	Pg	Pg	-	-
2	20	Pg	Pg+	Pg	Pg	-	Pg	Pg	Pg	-	-
2	TRIS	-	-	-	-	-	-	-	-	-	-
3	7	-	-	-	-	-	-	-	-	-	-
3	8	-	-	-	-	-	-	-	-	-	-
3	9	rP	-	-	-	-	Pg	Pg	rP	rP	Pg
3	20	Pg	Pg+	Pg	Pg	-	Pg	Pg	Pg	Pg	Pg
3	TRIS	-	-	-	-	-	-	-	-	-	-

## **Résumé:**

Les souches de *Xanthomonas citri* pv. *citri*, agent pathogène responsable du chancre asiatique des agrumes, sont subdivisées en deux pathotypes en fonction de leur gamme d'hôte: le pathotype A, à large gamme d'hôte, et le pathotype A\*, à gamme d'hôte restreinte. L'étude présentée ici vise à caractériser chez *Xcc* les TALEs (Transcription Activator Like Effectors), qui sont des protéines injectées dans la cellule hôte et pouvant jouer un rôle dans la restriction de la gamme d'hôte. Ces effecteurs de type III miment les facteurs de transcription eucaryotes pour activer certains gènes afin de favoriser l'infection. Les TALEs ont une structure particulière comprenant entre autre un domaine central constitué de répétitions de 34 acides aminés. Les travaux de Bogdanove et al. en 2010, ont abouti à la mise en évidence d'un code reliant certains acides aminés de la protéine (aux 12<sup>e</sup> et 13<sup>e</sup> positions de chaque répétition, le RVD) avec la séquence cible de l'ADN de la plante, dans la région promotrice du gène à activer. L'implication des TALEs dans la spécialisation de la gamme d'hôte a été mise en évidence à plusieurs reprises chez d'autres bactéries phytopathogènes ; Par ailleurs, un rôle dans le développement du chancre bactérien des agrumes a été clairement démontré. Du fait de leur structure en séquences répétées, les TALEs ont un potentiel évolutif fort, ce qui supporte l'hypothèse d'un rôle dans l'évolution du pouvoir pathogène proprement dit. Cette étude vise à comprendre leur implication dans la détermination du spectre d'hôte chez *Xcc*.

La caractérisation de la séquence des différents TALEs dans une collection mondiale de souches de *Xcc* a permis non seulement de confirmer une plus grande diversité de TALEs chez les souches de pathotype A, mais également de mettre en évidence la présence d'un homologue, *pthA3* fortement conservé chez ce même pathotype.

Enfin, ces travaux ont permis d'affiner la connaissance du TALE responsable de la formation des symptômes du chancre chez les différents pathotypes, TALE jusque là décrit comme composé de 17.5 répétitions (Al-Saadi *et al* en 2007). Les travaux de Southern blot, séquençage et mutagenèse ont confirmé que ce TALE est composé de 18.5 répétitions chez les souches du pathotype A\*.

**Mots clef :** *Xanthomonas citri* pv. *citri*, pathotypes, pouvoir pathogène, gamme d'hôte, TALE, RVD.

## **Abstract:**

*Xanthomonas citri* pv. *citri*, the causal agent of the Asiatic citrus canker, is divided in two pathotypes varying for of their host range. Pathotype A strains naturally infect a wide range of Citrus while the pathotype A\* strains have a narrow host range. Here we studied TALEs (transcription Activator Like Effector), which could play a key role in pathogenicity and host range restriction. These type III effectors mimic eukaryotic transcription factor that activate genes in order to enhance plant infection. TALEs have a specific structure with a central domain composed of 34 amino acids. Bogdanove *et al.* (2010) showed that a code exists between specific amino acids (12th and 13th position of each repetition) and the promotor of the target gene activated by these transcription factor. TALEs involvement in host range specialization has been clearly demonstrated in other phytopathogenic bacteria and their role in the development of the Asiatic citrus canker has already been shown. Moreover, the high evolutionary potential conferred by the repeat sequence structure of TALEs support the hypothesis of their role in the evolution of *Xcc* pathogenicity. This study aimed at understanding the role of TALEs in the specialization of *Xcc* host range.

We found that a high diversity of TALE content exists between strains of pathotype A and that the sequence of *pthA3* is highly conserved among pathotype A strains. We also showed by sequencing and mutagenesis that in A\* strains, the TALE involved in the formation of canker symptoms is composed of 18.5 repetitions and not 17.5 like in A strains and previously reported (Al-Saadi 2007)

**Keys words:** *Xanthomonas citri* pv. *citri*, pathotypes, pathogenicity, host range, TALE, RVD.